GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:40 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-32
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	29.4	20	6	Q9TR70	Q9tr70 ursus arcto
2	29	28.4	16	13	P82388	P82388 litoria ran
3	29	28.4	16	13	P82392	P82392 litoria ran
4	27	26.5	19	5	Q9TWK8	Q9twk8 trypanosoma
5	27	26.5	19	12	Q84274	Q84274 human papil
6	27	26.5	20	5	Q9TWV8	Q9twv8 dermatophag
7	26	25.5	19	10	O04817	O04817 sporobolus
8	26	25.5	20	2	Q9R4Z4	Q9r4z4 proteus mir
9	26	25.5	20	4	Q9UCA3	Q9uca3 homo sapien
10	25.5	25.0	20	10	Q9S8T0	Q9s8t0 artocarpus
11	25	24.5	18	5	Q9TWL4	Q9twl4 lucilia cup
12	25	24.5	19	2	Q06140	Q06140 neisseria m
13	25	24.5	19	2	Q9R5I1	Q9r5i1 brevibacter
14	24	23.5	18	5	Q9U8F7	Q9u8f7 drosophila
15	24	23.5	19	11	Q62637	Q62637 rattus norv
16	23	22.5	15	5	Q9TWF5	Q9twf5 artemia (br
17	23	22.5	16	13	P82389	P82389 litoria aur
18	23	22.5	16	13	P82390	P82390 litoria aur
19	23	22.5	17	4	Q9UCI3	Q9uci3 homo sapien
20	23	22.5	17	11	Q9QVC2	Q9qvc2 rattus sp.
21	23	22.5	18	6	Q9TT30	Q9tt30 saguinus la
22	23	22.5	18	10	Q9S892	Q9s892 glycine max
23	23	22.5	19	13	P82095	P82095 uperoleia m
24	23	22.5	20	4	Q9UD25	Q9ud25 homo sapien
25	22	21.6	14	11	Q10734	Q10734 mus musculu
26	22	21.6	15	12	Q88954	Q88954 vaccinia vi
27	22	21.6	16	2	Q9R514	Q9r514 porphyromon
28	22	21.6	17	6	Q9TRD5	Q9trd5 oryctolagus
29	22	21.6	18	4	Q14042	Q14042 homo sapien
30	22	21.6	19	2	Q9R5A9	Q9r5a9 arthrobacte
31	22	21.6	19	4	Q9UCG2	Q9ucg2 homo sapien
32	22	21.6	20	11	Q9QUZ7	Q9quz7 rattus sp.
33	22	21.6	20	13	Q9PRT8	Q9prt8 gadus morhu
34	22	21.6	20	13	Q9PRR3	Q9prr3 scyliorhinu
35	21.5	21.1	18	6	Q9TRB6	Q9trb6 sus scrofa
36	21.5	21.1	20	4	Q9UCM1	Q9ucm1 homo sapien
37	21	20.6	9	11	Q9QWT0	Q9qwt0 mus musculu
38	21	20.6	9	12	Q88953	Q88953 vaccinia vi
39	21	20.6	10	12	Q85462	Q85462 avian sarco
40	21	20.6	13	2	Q50038	Q50038 mycobacteri
41	21	20.6	15	11	Q9QUZ5	Q9quz5 mus sp. 38-
42	21	20.6	16	4	Q9UC48	Q9uc48 homo sapien
43	21	20.6	16	13	P82391	P82391 litoria aur
44	21	20.6	20	4	Q9UMU3	Q9umu3 homo sapien
45	21	20.6	20	6	Q9TRH6	Q9trh6 canis famil

ALIGNMENTS

RESULT 1
Q9TR70
ID Q9TR70 PRELIMINARY; PRT; 20 AA.
AC Q9TR70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HAPTOGLOBIN BETA CHAIN (FRAGMENT).
OS Ursus arctos (Brown bear) (Grizzly bear).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 95268704.
RA Mominoki K., Nakagawa-Tosa N., Morimatsu M., Syuto B., Saito M.;
RT "Haptoglobin in Carnivora: a unique molecular structure in bear, cat
RT and dog haptoglobins.";
RL Comp. Biochem. Physiol. Biochem. Mol. Biol. 110:785-789(1995).
SQ SEQUENCE 20 AA; 2020 MW; F605BCC82667DC3F CRC64;

Query Match 29.4%; Score 30; DB 6; Length 20;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy       9 GGFPSQAQV 17
           | || ||::
Db      10 GSFPXQAKM 18
```

RESULT 2
P82388
ID P82388 PRELIMINARY; PRT; 16 AA.
AC P82388;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE AUREIN 2.1/2.1.1.
OS Litoria raniformis, and Litoria aurea (Australian frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.

```
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: AUREIN 2.1 HAS ANTIMICROBIAL AND ANTICANCER ACTIVITY.
KW   Amidation; Antibiotic.
FT   PEPTIDE       3     16       AUREIN 2.1.1.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1616 MW;  1D9A5DADB4D240F9 CRC64;
                                  .

  Query Match             28.4%;  Score 29;  DB 13;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           :| ::||| | |
Db       4 DIVKKVVGAFGS 15


RESULT   3
P82392
ID   P82392      PRELIMINARY;      PRT;    16 AA.
AC   P82392;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AUREIN 2.5.
OS   Litoria raniformis, and Litoria aurea (Australian frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: HAS ANTIMICROBIAL AND ANTICANCER ACTIVITY.
KW   Amidation; Antibiotic.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1650 MW;  1D9A5DADB4DAE2F9 CRC64;


  Query Match             28.4%;  Score 29;  DB 13;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           :| ::||| | |
Db       4 DIVKKVVGAFGS 15


RESULT   4
Q9TWK8
ID   Q9TWK8      PRELIMINARY;      PRT;    19 AA.
AC   Q9TWK8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   HAPTOGLOBIN-RELATED PROTEIN BETA SUBUNIT (FRAGMENT).
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95232503.
RA   Smith A.B., Esko J.D., Hajduk S.L.;
RT   "Killing of trypanosomes by the human haptoglobin-related protein.";
RL   Science 268:284-286(1995).

SQ   SEQUENCE   19 AA;  2055 MW;  65BD135667C94056 CRC64;


  Query Match             26.5%;  Score 27;  DB 5;  Length 19;
  Best Local Similarity   55.6%;  Pred. No. 5.3e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQV 17
           | || ||::
Db      10 GSFPWQAKM 18


RESULT   5
Q84274
ID   Q84274      PRELIMINARY;      PRT;    19 AA.
AC   Q84274;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   E6 PROTEIN (FRAGMENT).
OS   Human papillomavirus type 25.
OC   Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillomavirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88089511.
RA   Krubke J., Kraus J., Delius H., Chow L., Broker T.R., Iftner T.,
RA   Pfister H.;
RT   "Genetic relationship among human papillomaviruses associated with
RT   benign and malignant tumours of patients with epidermodysplasia
RT   verruciformis.";
RL   J. Gen. Virol. 68:3091-3103(1987).
DR   EMBL; D00205; BAA00145.1; -.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1955 MW;  5C820487F69A5ED1 CRC64;


  Query Match             26.5%;  Score 27;  DB 12;  Length 19;
  Best Local Similarity   75.0%;  Pred. No. 5.3e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQV 17
           | | ||||
Db      11 GPPEQAQV 18


RESULT   6
Q9TWV8
ID   Q9TWV8      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWV8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   TRYPSIN-LIKE PROTEASE (DER F III ALLERGEN HOMOLOG) (FRAGMENT).
OS   Dermatophagoides farinae (House-dust mite).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC   Dermatophagoides.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93074427.
RA   Ando T., Homma R., Ino Y., Ito G., Miyahara A., Yamakawa H., Iwaki M.,
RA   Okumura Y., Suko M., Haida M.;
RT   "Is a trypsin-like protease of mites a Der f III allergen?";
RL   Jpn. J. Allergy 41:704-707(1992).
SQ   SEQUENCE   20 AA;  2087 MW;  961537685DB396A2 CRC64;


  Query Match             26.5%;  Score 27;  DB 5;  Length 20;
  Best Local Similarity   55.6%;  Pred. No. 5.7e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQA 15
           :|||  :||
```

```
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   FERREDOXIN--NAD(+) REDUCTASE (EC 1.18.1.3) (FRAGMENT).
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92041607.
RA   Ramachandra M., Seetharam R., Emptage M.H., Sariaslani F.S.;
RT   "Purification and characterization of a soybean flour-inducible
RT   ferredoxin reductase of Streptomyces griseus.";
RL   J. Bacteriol. 173:7106-7112(1991).
CC   -!- FUNCTION: COUPLE ELECTRON TRANSFER FROM NADH TO CYTOCHROME
CC       P450(SOY) IN THE PRESENCE OF FERREDOXIN.
CC   -!- CATALYTIC ACTIVITY: REDUCED FERREDOXIN + NAD(+) = OXIDIZED
CC       FERREDOXIN + NADH.
CC   -!- COFACTOR: FAD; REQUIRES MAGNESIUM.
DR   HSSP; P09063; 1LVL.
KW   Oxidoreductase; Flavoprotein; NAD; FAD; Magnesium.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1485 MW;  27D11A7C37AC0510 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 9.5e+02;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      7 VVGGFP 12
          ::|| |
Db      5 IIGGGP 10


RESULT  14
FIXA_RHILE
ID   FIXA_RHILE     STANDARD;      PRT;    18 AA.
AC   P14313;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   FIXA PROTEIN (FRAGMENT).
GN   FIXA.
OS   Rhizobium leguminosarum.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90066358.
RA   Hontelez J.G.J., Lankhorst R.K., Katinakis P., van den Bos R.C.,
RA   van Kammen A.;
RT   "Characterization and nucleotide sequence of a novel gene fixW
RT   upstream of the fixABC operon in Rhizobium leguminosarum.";
RL   Mol. Gen. Genet. 218:536-544(1989).
CC   -!- FUNCTION: MAY PLAY A ROLE IN A REDOX PROCESS INVOLVED IN NITROGEN
CC       FIXATION.
CC   -!- SUBUNIT: FIXA AND FIXB FORM A HETERODIMER (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X16521; CAA34527.1; -.
DR   PIR; PQ0022; PQ0022.
DR   INTERPRO; IPR000049; -.
DR   PROSITE; PS01065; ETF_BETA; PARTIAL.
KW   Electron transport; Nitrogen fixation.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2036 MW;  74973C8BA2087663 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
             |  ||: |
Db      11 PDSAQIRV 18


RESULT  15
TCBI_TRILO
ID   TCBI_TRILO     STANDARD;      PRT;    19 AA.
AC   P80070;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   TRICHOLONGIN BI AND BII.
OS   Trichoderma longibrachiatum.
OC   Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Trichoderma.
RN   [1]
RP   SEQUENCE, AND STRUCTURE BY NMR.
RX   MEDLINE; 92037622.
RA   Rebuffat S., Prigent Y., Auvin-Guette C., Bodo B.;
RT   "Tricholongins BI and BII, 19-residue peptaibols from Trichoderma
RT   longibrachiatum. Solution structure from two-dimensional NMR
RT   spectroscopy.";
RL   Eur. J. Biochem. 201:661-674(1991).
CC   -!- FUNCTION: TRICHOLONGINS ARE PEPTIDES WHICH DISPLAY ANTIFUNGAL
CC       AND ANTIBACTERIAL ACTIVITY. PROBABLY INTERACT WITH BIOLOGICAL
CC       MEMBRANES AND PROBABLY PRODUCES HOLES LEADING TO LEAKAGE.
CC   -!- PTM: AIB IS ALPHA-AMINO ISOBUTYRIC ACID.
CC   -!- PTM: IVA IS ISOVALINE.
KW   Antibiotic; Fungicide; Methylation; Acetylation; Hydroxylation;
KW   Membrane.
FT   MOD_RES       1      1       ACETYLATION.
FT   MOD_RES       1      1       METHYLATION (TO FORM AIB).
FT   MOD_RES       4      4       METHYLATION (TO FORM AIB).
FT   MOD_RES       5      5       METHYLATION (TO FORM AIB).
FT   MOD_RES       7      7       METHYLATION (TO FORM AIB).
FT   MOD_RES       8      8       METHYLATION (TO FORM AIB).
FT   MOD_RES       9      9       METHYLATION (TO FORM AIB).
FT   MOD_RES      12     12       METHYLATION (TO FORM AIB).
FT   MOD_RES      15     15       METHYLATION (TO FORM AIB).
FT   MOD_RES      16     16       METHYLATION (TO FORM AIB; IN BI).
FT   MOD_RES      16     16       ETHYLATION (TO FORM IVA; IN BII).
FT   MOD_RES      19     19       HYDROXYLATION.
SQ   SEQUENCE   19 AA;  1756 MW;  B908C2DFF883D238 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQVTV 19
             || :||  ::
Db       2 GFAAQAAASL 11


Search completed: December 21, 2000, 08:32:59
Job time: 442 sec
```

```
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "New antibiotic uperin peptides from the dorsal glands of the
RT   australian toadlet Uperoleia mjobergii.";
RL   Aust. J. Chem. 49:1325-1331(1996).
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1948; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   19 AA;  1949 MW;  24E4F83A6BA35F21 CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGG 10
           :| :::|||
Db       4 DIAKKLVGG 12


RESULT  10
TRYL_STREX
ID   TRYL_STREX     STANDARD;      PRT;     20 AA.
AC   P80420;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   TRYPSIN-LIKE PROTEASE (EC 3.4.21.-) (FRAGMENT).
OS   Streptomyces exfoliatus (Streptomyces hydrogenans).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=SMF13;
RX   MEDLINE; 95291424.
RA   Kim I.S., Lee K.J.;
RT   "Physiological roles of leupeptin and extracellular proteases in
RT   mycelium development of Streptomyces exfoliatus SMF13.";
RL   Microbiology 141:1017-1025(1995).
CC   -!- FUNCTION: INVOLVED IN MYCELIUM DIFFERENTIATION.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A; ALSO KNOWN AS THE
CC       ALPHA-LYTIC PROTEASE FAMILY.
DR   HSSP; P00775; 1SGT.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2129 MW;  4568F56D0E7393AF CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   36.4%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           | || |  :::
Db      10 GNFPFQQXLSM 20


RESULT  11
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;      9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83160932.
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   INTERPRO; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN       1      >9       PENTAXIN.
FT   NON_TER      9       9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQV 17
           ||| :: :
Db       1 GFPGKSLI 8


RESULT  12
UN46_CLOPA
ID   UN46_CLOPA     STANDARD;      PRT;     14 AA.
AC   P81362;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   UNKNOWN PROTEIN CP 46 FROM 2D-PAGE (FRAGMENT).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=W5;
RX   MEDLINE; 98291870.
RA   Flengsrud R., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.4, ITS MW IS: 38.2 KDA.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1550 MW;  198078F4C0367170 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 8.3e+02;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 IFQEVVG 9
           || :::|
Db       2 IFNDLIG 8


RESULT  13
FENR_STRGR
ID   FENR_STRGR     STANDARD;      PRT;     16 AA.
AC   P24134;
DT   01-MAR-1992 (Rel. 21, Created)
```

```
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 EIFQEVVGG 10
          :: ::||||
Db      4 DLAKKVVGG 12


RESULT   6
UP23_UPEIN
ID   UP23_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82029;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.3.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST B.CEREUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1974; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1975 MW;  2F34EF077BA35B70 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 EIFQEVVGG 10
          :: ::||||
Db      4 DLAKKVVGG 12


RESULT   7
COG1_CHIOP
ID   COG1_CHIOP     STANDARD;      PRT;    20 AA.
AC   P34153;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   COLLAGENOLYTIC PROTEASE 25 KDA II/III (EC 3.4.21.32) (FRAGMENT).
OS   Chionoecetes opilio (Crab-beetle).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Majoidea; Majidae; Chionoecetes.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEPATOPANCREAS;
RX   MEDLINE; 92120073.
RA   Klimova O.A., Vedishcheva Y.V., Strongin A.Y.;
RT   "Isolation and characteristics of collagenolytic enzymes from the
RT   hepatopancreas of the crab Chionoecetes opilio.";
RL   Dokl. Akad. Nauk SSSR 317:482-484(1991).
CC   -!- FUNCTION: THIS ENZYME IS A SERINE PROTEASE CAPABLE OF DEGRADING
CC       THE NATIVE TRIPLE HELIX OF COLLAGEN.
CC   -!- CATALYTIC ACTIVITY: BROAD SPECIFICITY; DEGRADES NATIVE COLLAGEN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   HSSP; P00771; 1AZZ.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Collagen degradation.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2204 MW;  CE0D7B996E7281A7 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   35.7%;  Pred. No. 5.5e+02;
  Matches    5;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      7 VVGGFPSQAQVTVH 20
          :|||  :    ||
Db      1 IVGGQEATPHTWVH 14


RESULT   8
CPAX_BOVIN
ID   CPAX_BOVIN     STANDARD;      PRT;    18 AA.
AC   P22779;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CYTOCHROME P450 2A (OLF2) (OLFACTIVE) (P52) (EC 1.14.14.1) (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 91027757.
RA   Lazard D., Tal N., Rubinstein M., Khen M., Lancet D., Zupko K.;
RT   "Identification and biochemical analysis of novel olfactory-specific
RT   cytochrome P-450IIA and UDP-glucuronosyl transferase.";
RL   Biochemistry 29:7433-7440(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A35704; A35704.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum; Olfaction.
FT   NON_TER       1      1
FT   VARIANT       6      6       G -> D.
FT   VARIANT      11     11       A -> E.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2058 MW;  F80746F76CCD77FF CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 QEIFQEVVG 9
          |: |:|: |
Db      9 QQAFKELQG 17


RESULT   9
UP27_UPEMJ
ID   UP27_UPEMJ     STANDARD;      PRT;    19 AA.
AC   P82039;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.7.
OS   Uperoleia mjobergii (Australian toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
RESULT   2
COCO_LIMPO
ID   COCO_LIMPO     STANDARD;      PRT;    14 AA.
AC   P35586;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   COCOONASE (EC 3.4.21.-) (FRAGMENT).
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 78037243.
RA   Law J.H., Dunn P.E., Kramer K.J.;
RT   "Insect proteases and peptidases.";
RL   Adv. Enzymol. Relat. Areas Mol. Biol. 45:389-425(1977).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   HSSP; P00760; 4TPI.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1452 MW;  1615FB1D73747570 CRC64;

  Query Match            24.5%;  Score 25;  DB 1;  Length 14;
  Best Local Similarity  80.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT   3
TRYP_FELCA
ID   TRYP_FELCA     STANDARD;      PRT;    16 AA.
AC   P81071;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT).
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PANCREAS;
RX   MEDLINE; 97235546.
RA   Steiner J.M., Medinger T.L., Williams D.A.;
RT   "Purification and partial characterization of feline trypsin.";
RL   Comp. Biochem. Physiol. 116B:87-93(1997).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT   PROPEP        1      8       ACTIVATION PEPTIDE.
FT   CHAIN         9    >16       TRYPSIN.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1825 MW;  A6D751BB58760A86 CRC64;

  Query Match            23.5%;  Score 24;  DB 1;  Length 16;
  Best Local Similarity  33.3%;  Pred. No. 4.4e+02;
  Matches    3;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 EVVGGFPSQ 14
           ::|||: ::
Db       8 KIVGGYTNR 16


RESULT   4
ALL7_OLEEU
ID   ALL7_OLEEU     STANDARD;      PRT;    19 AA.
AC   P81430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   POLLEN ALLERGEN OLE E 7 (OLE E VII) (FRAGMENT).
OS   Olea europaea (Common olive).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Lamiales; Oleaceae; Olea.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=POLLEN;
RA   Tejera M.L., Villalba M., Rodriguez R.;
RT   "Isolation and characterization of Ole e 7, a new allergen from olive
RT   tree pollen.";
RL   Submitted (JUL-1998) to the SWISS-PROT data bank.
KW   Allergen.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1986 MW;  999A5008C414A67E CRC64;

  Query Match            23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity  71.4%;  Pred. No. 5.2e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 PSQAQVT 18
           |||  ||
Db       2 PSQGTVT 8


RESULT   5
UP22_UPEIN
ID   UP22_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82028;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.2.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A WEAK ANTIBACTERIAL ACTIVITY AGAINST B.CEREUS,
CC       E.COLI, L.MESENTERIODES, L.INNOCUA, M.LUTEUS, P.HAEMOLYTICA,
CC       S.AUREUS AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1927 MW;  3283EF077BA35B70 CRC64;

  Query Match            23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity  44.4%;  Pred. No. 5.2e+02;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:59 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-32
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	25.5	15	1	GLN2_PINPS	P81107 pinus pinas
2	25	24.5	14	1	COCO_LIMPO	P35586 limulus pol
3	24	23.5	16	1	TRYP_FELCA	P81071 felis silve
4	24	23.5	19	1	ALL7_OLEEU	P81430 olea europa
5	24	23.5	19	1	UP22_UPEIN	P82028 uperoleia i
6	24	23.5	19	1	UP23_UPEIN	P82029 uperoleia i
7	24	23.5	20	1	COG1_CHIOP	P34153 chionoecete
8	23	22.5	18	1	CPAX_BOVIN	P22779 bos taurus
9	23	22.5	19	1	UP27_UPEMJ	P82039 uperoleia m
10	23	22.5	20	1	TRYL_STREX	P80420 streptomyce
11	22	21.6	9	1	SAMP_MUSCA	P19095 mustelus ca
12	22	21.6	14	1	UN46_CLOPA	P81362 clostridium
13	22	21.6	16	1	FENR_STRGR	P24134 streptomyce
14	22	21.6	18	1	FIXA_RHILE	P14313 rhizobium l
15	22	21.6	19	1	TCBI_TRILO	P80070 trichoderma
16	22	21.6	19	1	UP21_UPEIN	P82027 uperoleia i
17	21.5	21.1	20	1	COGC_PARCM	P20734 paralithode
18	21	20.6	17	1	UP37_UPEMJ	P82044 uperoleia m
19	21	20.6	18	1	AROF_STRRM	P80575 streptomyce
20	21	20.6	19	1	OXLA_OPHHA	P81383 ophiophagus
21	21	20.6	20	1	ATP4_SPIOL	P80085 spinacia ol
22	21	20.6	20	1	ELAS_GADMO	P32197 gadus morhu
23	21	20.6	20	1	TL18_SPIOL	P82536 spinacia ol
24	20	19.6	13	1	BPP1_BOTJA	P01020 bothrops ja
25	20	19.6	13	1	UN12_CLOPA	P81353 clostridium
26	20	19.6	15	1	CYSK_CLOPA	P81340 clostridium
27	20	19.6	15	1	MILT_ONCKE	P81037 oncorhynchu
28	20	19.6	15	1	UC25_MAIZE	P80631 zea mays (m
29	20	19.6	16	1	CT12_LITCI	P81840 litoria cit
30	20	19.6	16	1	LPK1_LOCMI	P20404 locusta mig
31	20	19.6	17	1	EFG_THEAQ	Q01697 thermus aqu
32	20	19.6	19	1	UP24_UPEIN	P82030 uperoleia i
33	20	19.6	20	1	ANCR_AGKBI	P33588 agkistrodon
34	19	18.6	10	1	FARP_MANSE	P18523 manduca sex
35	19	18.6	13	1	CRBL_VESTR	P17231 vespa tropi
36	19	18.6	15	1	ATP2_PINPS	P81663 pinus pinas
37	19	18.6	15	1	COXJ_THUOB	P80979 thunnus obe
38	19	18.6	15	1	ESTJ_MANSE	P19985 manduca sex
39	19	18.6	15	1	FIBA_ANAPL	P12801 anas platyr
40	19	18.6	16	1	CT11_LITCI	P81835 litoria cit
41	19	18.6	16	1	CT13_LITCI	P81846 litoria cit
42	19	18.6	19	1	HBB2_UROHA	P18992 uromastyx h
43	19	18.6	20	1	ALAT_PIG	P13191 sus scrofa
44	19	18.6	20	1	COG1_PARCM	P20731 paralithode
45	19	18.6	20	1	COG2_CHIOP	P34154 chionoecete

ALIGNMENTS

RESULT 1
GLN2_PINPS
ID GLN2_PINPS STANDARD; PRT; 15 AA.
AC P81107;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROBABLE GLUTAMINE SYNTHETASE LEAF ISOZYME (EC 6.3.1.2) (GLUTAMATE--
DE AMMONIA LIGASE) (S2205/S2287) (N47/N48) (FRAGMENT).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Pinus.
RN [1]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RA Plomion C., Costa P., Bahrman N., Frigerio J.-M.;
RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
RT dominant and codominant protein markers assayed on diploid tissue, in
RT a haploid-based genetic map.";
RL Silvae Genetica 46:161-165(1997).
RN [2]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RX MEDLINE; 99274088.
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -!- FUNCTION: THE LIGHT-MODULATED CHLOROPLAST ENZYME, ENCODED BY A
CC NUCLEAR GENE AND EXPRESSED PRIMARILY IN LEAVES, IS RESPONSIBLE FOR
CC THE REASSIMILATION OF THE AMMONIA GENERATED BY PHOTORESPIRATION
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -!- SUBUNIT: HOMOOCTAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.7, ITS MW IS: 42 KDA.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR INTERPRO; IPR001691; -.
DR PROSITE; PS00180; GLNA_1; PARTIAL.
DR PROSITE; PS00181; GLNA_ATP; PARTIAL.
KW Ligase; Chloroplast.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1614 MW; 24A2420BEFD60D27 CRC64;

Query Match 25.5%; Score 26; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VGGFP 12
|||:|
Db 10 VGGYP 14

```
Db      3 VSGGFAS 9


RESULT  14
PA0026
protein QA300027 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997
C;Accession: PA0026
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensiona
A;Reference number: PA0001
A;Accession: PA0026
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: leaf


  Query Match            22.5%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity  45.5%;  Pred. No. 1.1e+03;
  Matches    5;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      8 VGGFPSQAQVT 18
          | | |   |:|
Db      4 VYGXPFAXQIT 14


RESULT  15
A35704
cytochrome P450 olf2 - bovine (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 12-Mar-1999
C;Accession: A35704
R;Lazard, D.; Tal, N.; Rubinstein, M.; Khen, M.; Lancet, D.; Zupko, K.
Biochemistry 29, 7433-7440, 1990
A;Title: Identification and biochemical analysis of novel olfactory-specific cytochrome
A;Reference number: A35704; MUID:91027757
A;Accession: A35704
A;Molecule type: protein
A;Residues: 1-18 <LAZ>
C;Genetics:
A;Gene: CYP2A
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protei


  Query Match            22.5%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity  44.4%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 QEIFQEVVG 9
          |: |:|: |
Db      9 QQAFKELQG 17


Search completed: December 21, 2000, 08:30:07
Job time: 271 sec
```

```
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             24.5%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 VGGFPS 13
           |||| |
Db       5 VGGFAS 10


RESULT  9
A20190
hypodermin B - early cattle grub (fragment)
C;Species: Hypoderma lineatum (early cattle grub)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Jun-1993
C;Accession: A20190
R;Lecroisey, A.; Tong, N.T.; Keil, B.
Eur. J. Biochem. 134, 261-267, 1983
A;Title: Hypodermin B, a trypsin-related enzyme from the insect Hypoderma lineatum.
A;Reference number: A20190; MUID:83261874
A;Accession: A20190
A;Molecule type: protein
A;Residues: 1-16 <LEC>


  Query Match             24.5%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   80.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT  10
A45806
T-cell receptor beta chain C region type 1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Jul-1999
C;Accession: A45806
R;Dent, A.L.; Fink, P.J.; Hedrick, S.M.
J. Immunol. 143, 322-328, 1989
A;Title: Characterization of an alternative exon of the murine T cell receptor beta-chai
A;Reference number: A45806; MUID:89278666
A;Accession: A45806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-20 <DEN>
A;Cross-references: GB:M27225; NID:g339373; PIDN:AAA61099.1; PID:g553781
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: T-cell receptor


  Query Match             24.0%;  Score 24.5;  DB 2;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       2 EIFQEVVGGF-PSQAQV 17
           ::|   |  | ||:|::
Db       4 KVFPPEVAVFEPSEAEI 20


RESULT  11
I52698
hypothetical THRA1/BTR mutant fusion protein, cell line BT474 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C;Accession: I52698
R;Futreal, P.A.; Cochran, C.; Marks, J.R.; Iglehart, J.D.; Zimmerman, W.; Barrett, J.C.;
Cancer Res. 54, 1791-1794, 1994
A;Title: Mutation analysis of the THRA1 gene in breast cancer: deletion/fusion of the
A;Reference number: I52698; MUID:94185019
A;Accession: I52698
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <FUT>
A;Cross-references: GB:S71020; NID:g546111; PIDN:AAB30341.1; PID:g546112
C;Comment: This sequence is the chimeric product of a deletion or translocation mutat
C;Genetics:
A;Gene: THRA1/BTR
A;Map position: 17q11.2
C;Keywords: fusion protein


  Query Match             23.5%;  Score 24;  DB 4;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 7.5e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 FPSQAQVTVH 20
           |  | :| ||
Db       3 FSEQFRVQVH 12


RESULT  12
PC2241
heat shock protein 42A - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 17-Mar-1999
C;Accession: PC2241
R;Hatayama, T.; Yasuda, K.; Nishiyama, E.
Biochem. Biophys. Res. Commun. 204, 357-365, 1994
A;Title: Characterization of high-molecular-mass heat shock proteins and 42oC-specifi
A;Reference number: PC2238; MUID:95032120
A;Accession: PC2241
A;Molecule type: protein
A;Residues: 1-17 <HAT>
C;Keywords: heat shock; stress-induced protein


  Query Match             23.5%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   80.0%;  Pred. No. 8.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 EVVGG 10
           |:|||
Db       8 EIVGG 12


RESULT  13
PH1448
T-cell receptor alpha chain (clone A24/PEG2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1448
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.;
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompa
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1448
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             22.5%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 1e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 VVGGFPS 13
           | ||| |
```

```
RESULT   3
S56122
type I DNA methyltransferase M.EcoR124 I chain HsdM - Escherichia coli (fragments)
C;Species: Escherichia coli
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C;Accession: S56122
R;Webb, M.; Taylor, I.A.; Firman, K.; Kneale, G.G.
J. Mol. Biol. 250, 181-190, 1995
A;Title: Probing the domain structure of the type IC DNA methyltransferase M.EcoR124I by
A;Reference number: S56121; MUID:95333175
A;Accession: S56122
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <WEB>


  Query Match             25.5%;  Score 26;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GFPSQA 15
           |:||:|
Db       7 GYPSEA 12


RESULT   4
A61392
brain-associated small cell lung cancer antigen - human (fragment)
N;Alternate names: BASCA
C;Species: Homo sapiens (man)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C;Accession: A61392
R;Umezawa, Y.; Kuge, S.; Kikyo, N.; Shirai, T.; Watanabe, J.; Fujiwara, M.; Okabe, T.
Jpn. J. Clin. Oncol. 21, 251-255, 1991
A;Title: Identity of brain-associated small cell lung cancer antigen and the CD56 (NKH-1
A;Reference number: A61392; MUID:92046737
A;Accession: A61392
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <UME>


  Query Match             25.5%;  Score 26;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
           ||| :::|
Db       7 PSQGEISV 14


RESULT   5
C34917
catechol 1,2-dioxygenase (EC 1.13.11.1) beta chain - Pseudomonas sp. (fragment)
N;Alternate names: pyrocatechase beta chain
C;Species: Pseudomonas sp.
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 10-Feb-1995
C;Accession: C34917
R;Nakai, C.; Horiike, K.; Kuramitsu, S.; Kagamiyama, H.; Nozaki, M.
J. Biol. Chem. 265, 660-665, 1990
A;Title: Three isozymes of catechol 1,2-dioxygenase (pyrocatechase), alphaalpha, alphabe
A;Reference number: A34917; MUID:90110118
A;Accession: C34917
A;Molecule type: protein
A;Residues: 1-20 <NAK>
C;Keywords: heterodimer; homodimer; oxidoreductase


  Query Match             25.5%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVG 9
             |: |:|  |
Db      11 QQFFEEAAG 19


RESULT   6
S29636
jacalin beta-I chain - Artocarpus champeden (fragment)
C;Species: Artocarpus champeden
C;Date: 19-Mar-1997 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
C;Accession: S29636
R;Ngoc, L.D.; Brillard, M.; Hoebeke, J.
Biochim. Biophys. Acta 1156, 219-222, 1993
A;Title: The alpha- and beta-subunits of the jacalins are cleavage products from a 17
A;Reference number: S29635; MUID:93152601
A;Accession: S29636
A;Molecule type: protein
A;Residues: 1-20 <NGO>
A;Experimental source: seed
C;Complex: heterotetramer; two alpha and two beta chains
C;Function:
A;Description: seed storage protein
A;Note: lectin for D-galactosyl-beta-1->3-N-acetylgalactosamine, a tumor-associated T
C;Keywords: heterotetramer; lectin; seed; storage protein


  Query Match             25.0%;  Score 25.5;  DB 2;  Length 20;
  Best Local Similarity   56.2%;  Pred. No. 5.9e+02;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       3 IFQEVVGGFPSQAQVT 18
           | | |: | |  ||||
Db       6 ISQTVIVG-PWGAQVT 20


RESULT   7
B61168
cocoonase (EC 3.4.21.-) - Chinese oak silkmoth (fragment)
C;Species: Antheraea pernyi (Chinese oak silkmoth)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C;Accession: B61168
R;Kramer, K.J.; Felsted, R.L.; Law, J.H.
J. Biol. Chem. 248, 3021-3028, 1973
A;Title: Cocoonase. V. Structural studies on an insect serine protease.
A;Reference number: A61168; MUID:73166540
A;Accession: B61168
A;Molecule type: protein
A;Residues: 1-5 <KRA>
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-5/Product: cocoonase (fragment) #status experimental <MAT>


  Query Match             24.5%;  Score 25;  DB 2;  Length 5;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT   8
PH1436
T-cell receptor alpha chain (clone A24/PEF5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1436
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.;
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompa
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1436
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:07 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-32
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	28.4	16	2	A60551	leukocyte elastase
2	27	26.5	17	2	A61557	major merozoite su
3	26	25.5	12	2	S56122	type I DNA methylt
4	26	25.5	18	2	A61392	brain-associated s
5	26	25.5	20	2	C34917	catechol 1,2-dioxy
6	25.5	25.0	20	2	S29636	jacalin beta-I cha
7	25	24.5	5	2	B61168	cocoonase (EC 3.4.
8	25	24.5	15	2	PH1436	T-cell receptor al
9	25	24.5	16	2	A20190	hypodermin B - ear
10	24.5	24.0	20	2	A45806	T-cell receptor be
11	24	23.5	15	4	I52698	hypothetical THRA1
12	24	23.5	17	2	PC2241	heat shock protein
13	23	22.5	14	2	PH1448	T-cell receptor al
14	23	22.5	15	2	PA0026	protein QA300027 -
15	23	22.5	18	2	A35704	cytochrome P450 ol
16	23	22.5	20	2	S50022	trypsin-like prote
17	23	22.5	20	2	PU0033	aldose 1-epimerase
18	22.5	22.1	20	2	A56900	chymotrypsin I (EC
19	22	21.6	9	2	B20569	serum amyloid P-co
20	22	21.6	12	2	PH1183	T-cell receptor al
21	22	21.6	14	2	PH0747	T-cell receptor be
22	22	21.6	15	2	S26527	T-cell receptor al
23	22	21.6	16	2	C53113	Lys-gingipain form
24	22	21.6	16	2	A44897	ferredoxin--NADP+
25	22	21.6	17	2	B61334	trypsin (EC 3.4.21
26	22	21.6	18	2	PQ0022	fixA protein - Rhi
27	22	21.6	19	2	S29167	quinaldine oxidore
28	22	21.6	19	2	S74087	antibacterial prot
29	22	21.6	20	2	S50023	trypsin-like prote
30	22	21.6	20	2	S71601	recombination prot
31	21.5	21.1	20	2	S29108	glutathione transf
32	21.5	21.1	20	2	A61327	trypsin (EC 3.4.21
33	21.5	21.1	20	2	D34817	collagenolytic pro
34	21	20.6	12	2	PH1189	T-cell receptor al
35	21	20.6	12	2	PH1180	T-cell receptor al
36	21	20.6	12	2	PH1188	T-cell receptor al
37	21	20.6	12	2	PH1172	T-cell receptor al
38	21	20.6	12	2	PH1175	T-cell receptor al
39	21	20.6	13	2	A32734	enkephalin precurs
40	21	20.6	14	2	PA0015	seed storage prote
41	21	20.6	14	2	PS0278	ribulose-bisphosph
42	21	20.6	14	2	PH1450	T-cell receptor al
43	21	20.6	15	2	S26516	T-cell receptor al
44	21	20.6	15	2	S26524	T-cell receptor al
45	21	20.6	15	2	S26528	T-cell receptor al

ALIGNMENTS

RESULT 1
A60551
leukocyte elastase (EC 3.4.21.37) - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Mar-1999
C;Accession: A60551
R;Axelsson, L.; Bergenfeldt, M.; Bjoerk, P.; Olsson, R.; Ohlsson, K.
Scand. J. Clin. Lab. Invest. 50, 35-42, 1990
A;Title: Release of immunoreactive canine leukocyte elastase normally and in endotoxi
A;Reference number: A60551; MUID:90193608
A;Accession: A60551
A;Molecule type: protein
A;Residues: 1-16 <AXE>
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; leukocyte; lysosome; serine proteinase

Query Match 28.4%; Score 29; DB 2; Length 16;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      7 VVGGFPSQ 14
          :||| |:|
Db      1 IVGGRPAQ 8
```

RESULT 2
A61557
major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (strain Fc
C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: A61557
R;Heidrich, H.G.
Biol. Cell 64, 205-214, 1988
A;Title: Isolation and functional characterization of Plasmodium falciparum merozoite
A;Reference number: A61557; MUID:89150734
A;Accession: A61557
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-17 <HEI>
C;Keywords: surface antigen

Query Match 26.5%; Score 27; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      2 EIFQEVVGGFPS 13
          | :||:|   ||
Db      4 ESYQELVAVTPS 15
```

```
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,889
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  P38,659
;      REFERENCE/DOCKET NUMBER:  27866/32168
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-286-889-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
           :|||   ||
Db       1 VFQEXGAGF 9


RESULT  15
US-08-485-618-15
; Sequence 15, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  103
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,618
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32797
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-485-618-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
           :|||   ||
Db       1 VFQEXGAGF 9


Search completed: December 21, 2000, 08:31:50
Job time: 373 sec
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-666-473-114


  Query Match             27.5%;  Score 28;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
           |  |||||
Db       1 PHMAQVTV 8


RESULT  12
US-08-669-284B-26
; Sequence 26, Application US/08669284B
; Patent No. 5939534
;  GENERAL INFORMATION:
;    APPLICANT:  Inoue, Makoto
;    APPLICANT:  Kikuchi, Kaoru
;    APPLICANT:  Ishige, Yoko
;    APPLICANT:  Ito, Akira
;    APPLICANT:  Kimura, Toru
;    APPLICANT:  Nakayama, Chikao
;    APPLICANT:  No. 5939534uchi, Hiroshi
;    TITLE OF INVENTION:  NOVEL HUMAN CILIARY NEUROTROPHIC FACTORS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/669,284B
;      FILING DATE:  28-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/02269
;      FILING DATE:  27-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 06-268281
;      FILING DATE:  05-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 06-201504
;      FILING DATE:  02-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-350934
;      FILING DATE:  29-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-42041
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-669-284B-26


  Query Match             26.5%;  Score 27;  DB 2;  Length 8;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 FQEVVGG 10
            ||: |||
Db       2 FQQQVGG 8


RESULT  13
US-08-173-497-15
; Sequence 15, Application US/08173497
; Patent No. 5437958
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van Der Vieren, Monica
;    TITLE OF INVENTION:  No. 5437958el Human  2 Integrin Alpha
;    TITLE OF INVENTION:  Subunit
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 S. Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,497
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5437958and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/31363
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-173-497-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
            :|||   ||
Db       1 VFQEXGAGF 9


RESULT  14
US-08-286-889-15
; Sequence 15, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
```

```
;    TITLE OF INVENTION:  METHODS FOR THE TREATMENT OF
;    TITLE OF INVENTION:  GASTROINTESTINAL TRACT DISORDERS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/622,720A
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-396
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-622-720A-23


  Query Match             29.4%;  Score 30;  DB 2;  Length 17;
  Best Local Similarity   53.8%;  Pred. No. 21;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVTV 19
           |||| |  :  ||
Db       1 VVGGQPGNSPWTV 13


RESULT  10
US-08-943-363-112
; Sequence 112, Application US/08943363
; Patent No. 5837478
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5837478el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,363
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  112:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-363-112


  Query Match             28.4%;  Score 29;  DB 2;  Length 17;
  Best Local Similarity   75.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGG 10
             ||||  ||
Db      10 IFQEDAGG 17


RESULT  11
US-08-666-473-114
; Sequence 114, Application US/08666473
; Patent No. 5843713
;  GENERAL INFORMATION:
;    APPLICANT:  YOSHIDA, Aruto
;    APPLICANT:  TAKEUCHI, Makoto
;    TITLE OF INVENTION:  PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
;    TITLE OF INVENTION:  CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
;    TITLE OF INVENTION:  WITH MUCIN SUGAR CHAIN
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/666,473
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP95/02238
;      FILING DATE:  01-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-22101
;      FILING DATE:  09-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-269111
;      FILING DATE:  01-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/837
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  114:
```

```
;       STREET:  6300 Columbia Center, 701 Fifth Avenue
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  US
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/486,348A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C6
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-486-348A-41


  Query Match             30.4%;  Score 31;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   7
US-08-468-545B-41
; Sequence 41, Application US/08468545B
; Patent No. 5876712
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,545B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-468-545B-41


  Query Match             30.4%;  Score 31;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   8
US-08-466-680B-41
; Sequence 41, Application US/08466680B
; Patent No. 6075122
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,680B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-466-680B-41


  Query Match             30.4%;  Score 31;  DB 3;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   9
US-08-622-720A-23
; Sequence 23, Application US/08622720A
; Patent No. 5814308
;  GENERAL INFORMATION:
;    APPLICANT:  Zhang, Ke
```

```
Db      1 VVGGXPGNSPXTV 13


RESULT   3
5219991-8
;Patent No. 5219991
;     APPLICANT: LEONARD, EDWARD;SKEEL, ALISON H.;YOSHIMURA.
;TEIZO;APPELLA, ETTORE
;     TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN
;     NUMBER OF SEQUENCES: 12
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/586,085
;       FILING DATE: 21-SEP-1990
;SEQ ID NO:8:
;       LENGTH: 20
5219991-8


  Query Match             31.4%;  Score 32;  DB 5;  Length 20;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      7 VVGGFPSQAQVTV 19
          |||| |  :  ||
Db      1 VVGGHPGNSPWTV 13


RESULT   4
US-08-467-083-41
; Sequence 41, Application US/08467083
; Patent No. 5726023
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/NEU ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  68
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry
;       STREET:  6300 Columbia Center, 701 Fifth Avenue
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  US
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/467,083
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/414,417
;       FILING DATE:  06-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;       TELEX:  3723836 SEEDANBERRY
;  INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-467-083-41


  Query Match             30.4%;  Score 31;  DB 1;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      9 GGFPSQAQVTV 19
          |  |:|||: :
Db      1 GAMPNQAQMRI 11


RESULT   5
US-08-414-417B-41
; Sequence 41, Application US/08414417B
; Patent No. 5801005
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  69
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry LLP
;       STREET:  6300 Columbia Center, 701 Fifth Avenue
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  US
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/414,417B
;       FILING DATE:  31-MAR-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-414-417B-41


  Query Match             30.4%;  Score 31;  DB 1;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      9 GGFPSQAQVTV 19
          |  |:|||: :
Db      1 GAMPNQAQMRI 11


RESULT   6
US-08-486-348A-41
; Sequence 41, Application US/08486348A
; Patent No. 5846538
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  69
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry LLP
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:50 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-32
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	36	35.3	14	3	US-09-120-365-87	Sequence 87, Appl
2	34	33.3	19	5	5219991-12	Patent No. 5219991
3	32	31.4	20	5	5219991-8	Patent No. 5219991
4	31	30.4	15	1	US-08-467-083-41	Sequence 41, Appl
5	31	30.4	15	1	US-08-414-417B-41	Sequence 41, Appl
6	31	30.4	15	2	US-08-486-348A-41	Sequence 41, Appl
7	31	30.4	15	2	US-08-468-545B-41	Sequence 41, Appl
8	31	30.4	15	3	US-08-466-680B-41	Sequence 41, Appl
9	30	29.4	17	2	US-08-622-720A-23	Sequence 23, Appl
10	29	28.4	17	2	US-08-943-363-112	Sequence 112, App
11	28	27.5	12	2	US-08-666-473-114	Sequence 114, App
12	27	26.5	8	2	US-08-669-284B-26	Sequence 26, Appl
13	27	26.5	11	1	US-08-173-497-15	Sequence 15, Appl
14	27	26.5	11	1	US-08-286-889-15	Sequence 15, Appl
15	27	26.5	11	1	US-08-485-618-15	Sequence 15, Appl
16	27	26.5	11	1	US-08-362-652-15	Sequence 15, Appl
17	27	26.5	11	2	US-08-605-672-15	Sequence 15, Appl
18	27	26.5	11	2	US-08-482-293A-15	Sequence 15, Appl
19	27	26.5	11	2	US-08-943-363-15	Sequence 15, Appl
20	27	26.5	14	3	US-09-120-365-85	Sequence 85, Appl
21	27	26.5	16	2	US-08-474-696A-7	Sequence 7, Appli
22	27	26.5	17	3	US-08-256-747C-22	Sequence 22, Appl
23	27	26.5	19	2	US-08-474-696A-1	Sequence 1, Appli
24	27	26.5	19	2	US-08-474-696A-3	Sequence 3, Appli
25	27	26.5	20	1	US-08-221-816B-11	Sequence 11, Appl
26	26	25.5	6	1	US-07-890-422B-15	Sequence 15, Appl
27	26	25.5	10	1	US-07-890-422B-19	Sequence 19, Appl
28	26	25.5	11	1	US-08-081-539-106	Sequence 106, App
29	26	25.5	11	1	US-08-466-647-106	Sequence 106, App
30	26	25.5	13	3	US-08-834-314-5	Sequence 5, Appli
31	26	25.5	14	3	US-09-120-365-82	Sequence 82, Appl
32	26	25.5	14	3	US-09-120-365-83	Sequence 83, Appl
33	26	25.5	14	3	US-09-120-365-84	Sequence 84, Appl
34	26	25.5	16	2	US-08-480-190-175	Sequence 175, App
35	26	25.5	16	2	US-08-488-379-175	Sequence 175, App
36	26	25.5	16	4	PCT-US93-07545-175	Sequence 175, App
37	25.5	25.0	20	1	US-08-218-025A-63	Sequence 63, Appl
38	25	24.5	5	3	US-08-570-761-3	Sequence 3, Appli
39	25	24.5	6	1	US-07-890-422B-14	Sequence 14, Appl
40	25	24.5	6	3	US-08-570-761-2	Sequence 2, Appli
41	25	24.5	7	1	US-08-482-880-32	Sequence 32, Appl
42	25	24.5	7	2	US-08-273-274-32	Sequence 32, Appl
43	25	24.5	7	2	US-08-475-041-32	Sequence 32, Appl
44	25	24.5	7	2	US-08-484-773-32	Sequence 32, Appl
45	25	24.5	9	2	US-08-934-222-89	Sequence 89, Appl

ALIGNMENTS

```
RESULT 1
US-09-120-365-87
; Sequence 87, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/120,365
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Nicotiana
US-09-120-365-87


  Query Match             35.3%;  Score 36;  DB 3;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 1.3;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           ||  ||| | ||
Db       3 EIEDEVVAGLPS 14


RESULT 2
5219991-12
;Patent No. 5219991
;    APPLICANT: LEONARD, EDWARD;SKEEL, ALISON H.;YOSHIMURA.
;TEIZO;APPELLA, ETTORE
;    TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN
;    NUMBER OF SEQUENCES: 12
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/586,085
;      FILING DATE: 21-SEP-1990
;SEQ ID NO:12:
;      LENGTH: 19
5219991-12


  Query Match             33.3%;  Score 34;  DB 5;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 4.4;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVTV 19
           |||| |  :  ||
```

```
DE   NF2 PROTEIN (FRAGMENT).
GN   NF2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94348501.
RA   Arakawa H., Hayashi N., Nagase H., Ogawa M., Nakamura Y.;
RT   "Alternative splicing of the NF2 gene and its mutation analysis of
RT   breast and colorectal cancers.";
RL   Hum. Mol. Genet. 3:565-568(1994).
DR   EMBL; S73854; AAB31737.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   17 AA;  1960 MW;  21B8200F8BE9653E CRC64;


  Query Match             22.9%;  Score 24;  DB 4;  Length 17;
  Best Local Similarity   30.8%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       4 TFRFPRPDGREAV 16
           | : |:  ||  :
Db       3 TIKKPQAQGRRPI 15


RESULT  15
Q9TRH5
ID   Q9TRH5      PRELIMINARY;      PRT;    17 AA.
AC   Q9TRH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ALPHA-S1-CASEIN HOMOLOG (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93231344.
RA   Neuteboom B., Giuffrida M.G., Conti A.;
RT   "Isolation of a new ligand-carrying casein fragment from bovine
RT   mammary gland microsomes.";
RL   FEBS Lett. 305:189-191(1992).
SQ   SEQUENCE   17 AA;  1950 MW;  8890EF0685A0537F CRC64;


  Query Match             22.9%;  Score 24;  DB 6;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREAV 16
           ||   |:| |
Db       5 FPEVFGKEKV 14


Search completed: December 21, 2000, 08:35:34
Job time: 597 sec
```

```
Qy      7 FPRPDGREA 15
          |||  | ||
Db      6 FPRIYGHEA 14


RESULT  10
O06946
ID   O06946      PRELIMINARY;      PRT;    17 AA.
AC   O06946;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   HYPOTHETICAL PROTEIN (ORF238) (FRAGMENT).
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE; 98036037.
RA   Hayes F., Lubetzki S.A., Sherratt D.J.;
RT   "Salmonella typhimurium specifies a circular chromosome dimer
RT   resolution system which is homologous to the Xer site-specific
RT   recombination system of Escherichia coli.";
RL   Gene 198:105-110(1997).
DR   EMBL; U92525; AAC45779.1; -.
KW   Hypothetical protein.
FT   NON_TER      17      17
SQ   SEQUENCE   17 AA;  2072 MW;  BEB65CE8F0F9F529 CRC64;


  Query Match             25.7%;  Score 27;  DB 2;  Length 17;
  Best Local Similarity   75.0%;  Pred. No. 5.2e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 RFPRPDGR 13
          || || ||
Db      2 RFYRPLGR 9


RESULT  11
Q9UC91
ID   Q9UC91      PRELIMINARY;      PRT;    17 AA.
AC   Q9UC91;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ONCOFETAL-LAMININ BINDING COLLAGEN ALPHA 1(III)CHAIN, OF-LB COLLAGEN
DE   ALPHA 1(III).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95169134.
RA   Minafra I.P., Andriolo M., Basirico L., Aquino A., Minafra S.,
RA   Boutillon M.M., van der Rest M.;
RT   "Onco-fetal/laminin-binding collagen from colon carcinoma: detection
RT   of new sequences.";
RL   Biochem. Biophys. Res. Commun. 207:852-859(1995).
SQ   SEQUENCE   17 AA;  1710 MW;  E6805580DA20B55E CRC64;


  Query Match             23.8%;  Score 25;  DB 4;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      7 FPRPDGRE 14
          || | | :
Db      2 FPGPKGND 9


RESULT  12
```

```
Q9TWL3
ID   Q9TWL3      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWL3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   INITIATORIN (FRAGMENT).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95218779.
RA   Aigaki T., Kasuga H., Nagaoka S., Osanai M.;
RT   "Purification and partial amino acid sequence of initiatorin, a
RT   prostatic endopeptidase of the silkworm, Bombyx mori.";
RL   Insect Biochem. Mol. Biol. 24:969-975(1994).
DR   INTERPRO; IPR001254; -.
DR   PFAM; PF00089; trypsin; 1.
SQ   SEQUENCE   20 AA;  2205 MW;  8BE047E96CFFF1BA CRC64;


  Query Match             23.8%;  Score 25;  DB 5;  Length 20;
  Best Local Similarity   55.6%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     12 GREAVAYRF 20
          || || : |
Db      4 GRRAVPHSF 12


RESULT  13
Q14001
ID   Q14001      PRELIMINARY;      PRT;    17 AA.
AC   Q14001;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-1997 (TrEMBLrel. 03, Last annotation update)
DE   CYCLIC NUCLEOTIDE PHOSPHODIESTERASE (FRAGMENT).
GN   CGIPDE1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 97079687.
RA   Loebbert R.W., Winterpacht A., Seipel B., Zabel B.U.;
RT   "Molecular cloning and chromosomal assignment of the human homologue
RT   of the rat cGMP-inhibited phosphodiesterase 1 (PDE3A)--a gene involved
RT   in fat metabolism located at 11p 15.1.";
RL   Genomics 37:211-218(1996).
DR   EMBL; X95522; CAA64776.1; -.
FT   NON_TER      17      17
SQ   SEQUENCE   17 AA;  2057 MW;  69D117C4FAE11540 CRC64;


  Query Match             22.9%;  Score 24;  DB 4;  Length 17;
  Best Local Similarity   54.5%;  Pred. No. 1.6e+03;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      9 RPDGREAVAYR 19
          | | |:| | |
Db      2 RRDERDAKAMR 12


RESULT  14
Q16231
ID   Q16231      PRELIMINARY;      PRT;    17 AA.
AC   Q16231;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
```

```
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria iinumae.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRA 377;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000218; AAC36545.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1713 MW;  CEF73B5B28A4D952 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIYGHEA 14


RESULT   7
O82406
ID   O82406      PRELIMINARY;      PRT;    16 AA.
AC   O82406;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria moschata.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRA 157;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000219; AAC36546.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1713 MW;  CEF73B5B28A4D952 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIYGHEA 14


RESULT   8
O82407
ID   O82407      PRELIMINARY;      PRT;    16 AA.
AC   O82407;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria viridis.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRA 341;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000220; AAC36547.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1682 MW;  CEF73B5B28A4D817 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIYGHEA 14


RESULT   9
O82781
ID   O82781      PRELIMINARY;      PRT;    16 AA.
AC   O82781;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria vesca.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=YELLOW WONDER, BARON SOLEMACHER (BS);
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000216; AAC36543.1; -.
DR   EMBL; AF000214; AAC36541.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1682 MW;  CEF73B5B28A4D817 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 87213162.
RA   Koiwai O., Kaneda T., Morishita R.;
RT   "Analysis of human terminal deoxynucleotidyl transferase cDNA
RT   expressible in mammalian cells.";
RL   Biochem. Biophys. Res. Commun. 144:185-190(1987).
DR   EMBL; M26144; AAA74588.1; -.
KW   Transferase.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1442 MW;  25B7D365F34FC408 CRC64;


  Query Match             25.7%;  Score 27;  DB 4;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 RPDGRE 14
           :||||:
Db       3 QPDGRQ 8


RESULT   3
O82402
ID   O82402      PRELIMINARY;      PRT;    16 AA.
AC   O82402;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria nubicola.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRA520;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000213; AAC36540.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1682 MW;  CEF73B5B28A4D817 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIYGHEA 14


RESULT   4
O82403
ID   O82403      PRELIMINARY;      PRT;    16 AA.
AC   O82403;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria vesca.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PAWTUCKAWAY;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000215; AAC36542.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1666 MW;  CEF73B5B28A4CA87 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIFGHEA 14


RESULT   5
O82404
ID   O82404      PRELIMINARY;      PRT;    16 AA.
AC   O82404;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ALCOHOL DEHYDROGENASE (EC 1.1.1.1) (ALDEHYDE REDUCTASE) (FRAGMENT).
GN   ADH.
OS   Fragaria nilgerrensis.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Rosaceae;
OC   Fragaria.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERLIN 1;
RA   Yu H., Davis T.M.;
RT   "Genetic relationships among Fragaria species based on RAPDs and an
RT   alcohol dehydrogenase (ADH) gene.";
RL   Genome 0:0-0(1997).
CC   -!- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE +
CC       NADH.
CC   -!- COFACTOR: ZINC OR IRON.
DR   EMBL; AF000217; AAC36544.1; -.
KW   Oxidoreductase.
FT   NON_TER       1      1
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1713 MW;  CEF73B5B28A4D952 CRC64;


  Query Match             25.7%;  Score 27;  DB 10;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 4.9e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREA 15
           |||  | ||
Db       6 FPRIYGHEA 14


RESULT   6
O82405
ID   O82405      PRELIMINARY;      PRT;    16 AA.
AC   O82405;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:33 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-3
Perfect score: 105
Sequence: 1 VAVTFRFPRPDGREAVAYRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	32.4	19	2	Q9R517	Q9r517 mycobacteri
2	27	25.7	13	4	Q14182	Q14182 homo sapien
3	27	25.7	16	10	O82402	O82402 fragaria nu
4	27	25.7	16	10	O82403	O82403 fragaria ve
5	27	25.7	16	10	O82404	O82404 fragaria ni
6	27	25.7	16	10	O82405	O82405 fragaria ii
7	27	25.7	16	10	O82406	O82406 fragaria mo
8	27	25.7	16	10	O82407	O82407 fragaria vi
9	27	25.7	16	10	O82781	O82781 fragaria ve
10	27	25.7	17	2	O06946	O06946 salmonella
11	25	23.8	17	4	Q9UC91	Q9uc91 homo sapien
12	25	23.8	20	5	Q9TWL3	Q9tw13 bombyx mori
13	24	22.9	17	4	Q14001	Q14001 homo sapien
14	24	22.9	17	4	Q16231	Q16231 homo sapien
15	24	22.9	17	6	Q9TRH5	Q9trh5 bos taurus
16	24	22.9	17	12	Q78328	Q78328 human immun
17	24	22.9	19	2	Q9R4J9	Q9r4j9 mycobacteri
18	23	21.9	12	8	Q31851	Q31851 arabidopsis
19	23	21.9	15	2	Q9R5L9	Q9r519 bacillus st
20	23	21.9	15	2	Q9R5L8	Q9r518 bacillus th
21	23	21.9	15	4	O00604	O00604 homo sapien
22	23	21.9	19	6	Q9TRH3	Q9trh3 sus scrofa
23	22	21.0	17	6	Q9TQZ8	Q9tqz8 sus scrofa
24	22	21.0	18	12	Q87589	Q87589 chimpanzee
25	22	21.0	18	12	Q87591	Q87591 chimpanzee
26	22	21.0	18	12	Q87593	Q87593 chimpanzee
27	22	21.0	20	3	P82263	P82263 aspergillus
28	22	21.0	20	6	Q9TRS2	Q9trs2 sus scrofa
29	21	20.0	10	11	Q9QVJ7	Q9qvj7 mus sp. mep
30	21	20.0	11	7	O77899	O77899 oreochromis
31	21	20.0	11	7	O77900	O77900 oreochromis
32	21	20.0	11	7	O77901	O77901 oreochromis
33	21	20.0	11	7	O77902	O77902 oreochromis
34	21	20.0	11	7	O77903	O77903 oreochromis
35	21	20.0	11	7	O77904	O77904 oreochromis
36	21	20.0	11	7	O77905	O77905 oreochromis
37	21	20.0	11	7	O77916	O77916 oreochromis
38	21	20.0	11	7	O77917	O77917 oreochromis
39	21	20.0	11	7	O77921	O77921 pseudotroph
40	21	20.0	12	2	Q50959	Q50959 neisseria g
41	21	20.0	13	4	Q9UM84	Q9um84 homo sapien
42	21	20.0	13	4	Q9UEE2	Q9uee2 homo sapien
43	21	20.0	13	11	Q63940	Q63940 rattus norv
44	21	20.0	14	2	Q52636	Q52636 escherichia
45	21	20.0	15	2	Q52304	Q52304 escherichia

ALIGNMENTS

RESULT 1
Q9R517
ID Q9R517 PRELIMINARY; PRT; 19 AA.
AC Q9R517;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE ISONICOTINATE DEHYDROGENASE (FRAGMENT).
OS Mycobacterium.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae.
RN [1]
RP SEQUENCE.
RX MEDLINE; 94103749.
RA Kretzer A., Frunzke K., Andreesen J.R.;
RT "Catabolism of isonicotinate by Mycobacterium sp. INA1: extended
RT description of the pathway and purification of the molybdoenzyme
RT isonicotinate dehydrogenase.";
RL J. Gen. Microbiol. 139:2763-2772(1993).
SQ SEQUENCE 19 AA; 2215 MW; 6A4FB97459902F0B CRC64;

Query Match 32.4%; Score 34; DB 2; Length 19;
Best Local Similarity 58.3%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       5 FRFPRPDGREAV 16
           :||   ||| ||
Db       6 YRFAEDDGRIAV 17
```

RESULT 2
Q14182
ID Q14182 PRELIMINARY; PRT; 13 AA.
AC Q14182;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE DEOXYNUCLEOTIDYLTRANSFERASE (FRAGMENT).
GN DNTT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Job time: 437 sec

```
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   INTERPRO; IPR002181; -.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma; Sulfatation.
FT   MOD_RES       3      3       SULFATATION.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2296 MW;  921A2B02D5F6891D CRC64;


  Query Match             21.0%;  Score 22;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     11 DGREAVAY 18
          |||  | :
Db      9 DGRTKVTF 16


RESULT  13
LIGA_TRAVE
ID   LIGA_TRAVE     STANDARD;      PRT;    13 AA.
AC   P20011;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   LIGNINASE A (EC 1.11.1.-) (LIGNIN PEROXIDASE) (FRAGMENT).
OS   Trametes versicolor (White-rot fungus).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Aphyllophorales;
OC   Coriolaceae; Trametes.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89211432.
RA   Joensson L., Karlsson O., Lundquist K., Nyman P.O.;
RT   "Trametes versicolor ligninase: isozyme sequence homology and
RT   substrate specificity.";
RL   FEBS Lett. 247:143-146(1989).
CC   -!- CATALYTIC ACTIVITY: DEPOLYMERIZATION OF LIGNIN. CATALYSES THE
CC       C(ALPHA)-C(BETA) CLEAVAGE OF THE PROPYL SIDE CHAINS OF LIGNIN.
CC   -!- PATHWAY: FIRST STEP IN LIGNIN DEGRADATION.
CC   -!- SIMILARITY: BELONGS TO THE PEROXIDASE FAMILY. LIGNINASE SUBFAMILY.
DR   PIR; S04013; S04013.
DR   INTERPRO; IPR002016; -.
DR   PROSITE; PS00435; PEROXIDASE_1; PARTIAL.
DR   PROSITE; PS00436; PEROXIDASE_2; PARTIAL.
KW   Oxidoreductase; Peroxidase; Heme; Glycoprotein; Multigene family;
KW   Lignin degradation.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1298 MW;  22C50ED5872A4338 CRC64;


  Query Match             20.0%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 9.9e+02;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     10 PDGR 13
          |||:
Db      4 PDGK 7


RESULT  14
MK2A_PALPR
ID   MK2A_PALPR     STANDARD;      PRT;    15 AA.
AC   P80409;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   METALNIKOWIN IIA.
OS   Palomena prasina.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Paraneoptera; Hemiptera; Euhemiptera;
OC   Heteroptera; Panheteroptera; Pentatomomorpha; Pentatomoidea;
OC   Pentatomidae; Palomena.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEMOLYMPH;
RA   Chernysh S., Cociancich S., Briand J.-P., Hetru C., Bulet P.;
RT   "The inducible antibacterial peptides of the hemipteran insect
RT   Palomena prasina. Identification of a unique family of proline-rich
RT   peptides and of a novel insect defensin.";
RL   J. Insect Physiol. 42:81-89(1996).
CC   -!- FUNCTION: ANTIBACTERIAL PEPTIDE ACTIVE AGAINST GRAM-NEGATIVE
CC       BACTERIA.
CC   -!- INDUCTION: BY BACTERIAL INFECTION.
KW   Antibiotic; Insect immunity.
SQ   SEQUENCE   15 AA;  1893 MW;  23835D063B946299 CRC64;


  Query Match             20.0%;  Score 21;  DB 1;  Length 15;
  Best Local Similarity   60.0%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FPRPD 11
          :|||:
Db     11 WPRPN 15


RESULT  15
SODM_STRGR
ID   SODM_STRGR     STANDARD;      PRT;    15 AA.
AC   P80733;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SUPEROXIDE DISMUTASE [MN] (EC 1.15.1.1) (FRAGMENT).
GN   SOD2.
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=KCTC 9006;
RX   MEDLINE; 97056064.
RA   Youn H.-D., Youn H., Lee J.-W., Yim Y.-I., Lee J.K., Hah Y.C.,
RA   Kang S.-O.;
RT   "Unique isozymes of superoxide dismutase in Streptomyces griseus.";
RL   Arch. Biochem. Biophys. 334:341-348(1996).
CC   -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC       CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC   -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC   -!- SUBUNIT: TETRAMER (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC       FAMILY.
DR   HSSP; P80293; 1AR4.
DR   INTERPRO; IPR001189; -.
DR   PFAM; PF00081; sodfe; 1.
DR   PROSITE; PS00088; SOD_MN; PARTIAL.
KW   Oxidoreductase; Manganese.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1685 MW;  327993F710861372 CRC64;


  Query Match             20.0%;  Score 21;  DB 1;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 TFRFPRP 10
          |:  | |
Db      2 TYTLPEP 8


Search completed: December 21, 2000, 08:32:54
```

```
PHSL_DESBN
ID   PHSL_DESBN     STANDARD;      PRT;    19 AA.
AC   P13066;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PERIPLASMIC [NIFESE] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESE
DE   HYDROGENLYASE LARGE CHAIN) (FRAGMENT).
OS   Desulfovibrio baculatus (strain Norway 4).
OC   Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88106446.
RA   Prickril B.C., He S.H., Li C., Menon N., Choi E.S., Przybyla A.E.,
RA   Dervartanian D.V., Peck H.D. Jr., Fauque G., le Gall J., Teixeira M.,
RA   Moura I., Moura J.J.G., Patil D., Huynh B.H.;
RT   "Identification of three classes of hydrogenase in the genus,
RT   Desulfovibrio.";
RL   Biochem. Biophys. Res. Commun. 149:369-377(1987).
CC   -!- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) = 2 OXIDIZED
CC       FERREDOXIN + H(2).
CC   -!- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE
CC       SELENOCYSTEINE.
CC   -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE
CC       SUBUNIT FAMILY.
DR   PIR; H27480; H27480.
DR   INTERPRO; IPR001501; -.
DR   PROSITE; PS00507; NI_HGENASE_L_1; PARTIAL.
DR   PROSITE; PS00508; NI_HGENASE_L_2; PARTIAL.
KW   Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1942 MW;  2BFCD2D360F00367 CRC64;


  Query Match             21.9%;  Score 23;  DB 1;  Length 19;
  Best Local Similarity   30.0%;  Pred. No. 7.1e+02;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      8 PRPDGREAVA 17
          |  ||:: ::
Db      6 PAADGKKKIS 15


RESULT  10
UXA6_CHLTR
ID   UXA6_CHLTR     STANDARD;      PRT;    10 AA.
AC   P38007;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE FROM ELEMENTARY BODY (FRAGMENT).
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=L2/434/BU;
RA   Bini L., Santucci A., Magi B., Marzocchi B., Sanchez-Campillo M.,
RA   Comanducci M., Christianen G., Birkelund S., Vtretou E., Ratti G.,
RA   Pallini V.;
RL   Submitted (SEP-1994) to the SWISS-PROT data bank.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 6.44, ITS MW IS: 38.6 KDA.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1243 MW;  DAD39A33304B5339 CRC64;


  Query Match             21.0%;  Score 22;  DB 1;  Length 10;
  Best Local Similarity   30.0%;  Pred. No. 5.2e+02;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 VTFRFPRPDG 12


          : |:: : ||
Db      1 MNFKYIKKDG 10


RESULT  11
TKN1_KASMA
ID   TKN1_KASMA     STANDARD;      PRT;    12 AA.
AC   P08613;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYLAMBATES KASSININ (GLU(2)-PRO(5) KASSININ).
OS   Kassina maculata (African rhacophorid frog) (Hylambates maculatus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;
OC   Kassina.
RN   [1]
RP   SEQUENCE.
RA   Yasuhara T., Nakajima T., Erspamer G.F., Erspamer V.;
RT   "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and
RT   hylambatin, in the skin of the African rhacophorid frog Hylambates
RT   maculatus.";
RL   Biomed. Res. 2:613-617(1981).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S10059; S10059.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Amphibian skin.
FT   MOD_RES      12     12       AMIDATION.
SQ   SEQUENCE   12 AA;  1376 MW;  3E756D279DD6DAB7 CRC64;


  Query Match             21.0%;  Score 22;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 6.3e+02;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 PRPD 11
          |:||
Db      3 PKPD 6


RESULT  12
FIBB_HORSE
ID   FIBB_HORSE     STANDARD;      PRT;    19 AA.
AC   P14471;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   FIBRINOPEPTIDE B.
OS   Equus caballus (Horse), and Equus asinus (Donkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=HORSE;
RA   Blomback B., Blomback M., Grondahl N.J.;
RT   "Studies on fibrinopeptides from mammals.";
RL   Acta Chem. Scand. 19:1789-1791(1965).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=E.ASINUS;
RA   Blomback B., Blomback M., Grondahl N.J., Holmberg E.;
RT   "Structure of fibrinopeptides-its relation to enzyme specificity and
RT   phylogeny and classification of species.";
RL   Ark. Kemi 25:411-428(1966).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
```

```
  Best Local Similarity   50.0%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 DGREAVAYRF 20
           || :||   |
Db       2 DGSDAVGGEF 11


RESULT   6
CAQS_RAT
ID   CAQS_RAT       STANDARD;      PRT;    20 AA.
AC   P19633;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CALSEQUESTRIN, SKELETAL MUSCLE ISOFORM (ASPARTACTIN) (LAMININ-BINDING
DE   PROTEIN) (FRAGMENT).
GN   CASQ1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88331073.
RA   Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT   "Isolation and characterization of a laminin-binding protein from rat
RT   and chick muscle.";
RL   J. Cell Biol. 107:687-697(1988).
CC   -!- FUNCTION: CALSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC       CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC       IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH
CC       A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC       TO 50 MOLES OF CALCIUM. ALSO BINDS LAMININ.
CC   -!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE
CC       SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF
CC       FAST SKELETAL MUSCLE CELLS. ASPARTACTIN IS FOUND IN THE BASAL
CC       LAMINA SURROUNDING INDIVIDUAL MUSCLE FIBERS.
CC   -!- TISSUE SPECIFICITY: SKELETAL AND HEART MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
DR   PIR; A31049; A31049.
DR   INTERPRO; IPR001393; -.
DR   PROSITE; PS00864; CALSEQUESTRIN_2; PARTIAL.
DR   PROSITE; PS00863; CALSEQUESTRIN_1; 1.
KW   Muscle; Glycoprotein; Calcium-binding.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2238 MW;  92ADE04FC2A69280 CRC64;


  Query Match             22.9%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FPRPDGREAV 16
           ||  || : |
Db       6 FPEYDGVDRV 15


RESULT   7
CEP1_ACHFU
ID   CEP1_ACHFU     STANDARD;      PRT;    11 AA.
AC   P22790;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARDIO-EXCITATORY PEPTIDE-1 (ACEP-1).
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Achatinacea; Achatinidae; Achatina.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=FERUSSAC; TISSUE=HEART ATRIUM;
RX   MEDLINE; 90211261.
RA   Fujimoto K., Ohta N., Yoshida M., Kubota I., Muneoka Y., Kobayashi M.;
RT   "A novel cardio-excitatory peptide isolated from the atria of the
RT   African giant snail, Achatina fulica.";
RL   Biochem. Biophys. Res. Commun. 167:777-783(1990).
CC   -!- FUNCTION: POTENTIATES THE BEAT OF THE VENTRICLE, AND HAS ALSO
CC       EXCITATORY ACTIONS ON THE PENIS RETRACTOR MUSCLE, THE BUCCAL
CC       MUSCLE AND THE IDENTIFIED NEURONS CONTROLLING THE BUCCAL MUSCLE
CC       MOVEMENT OF ACHATINA.
CC   -!- SIMILARITY: TO POSSIBLE PEPTIDE L5 FROM APLYSIA.
DR   PIR; A34662; A34662.
KW   Hormone; Amidation.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1305 MW;  82D6D5B9C7741365 CRC64;


  Query Match             21.9%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 RPDGR 13
           || ||
Db       6 RPQGR 10


RESULT   8
ATPB_CANFA
ID   ATPB_CANFA     STANDARD;      PRT;    19 AA.
AC   P99504;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   ATP SYNTHASE BETA CHAIN, MITOCHONDRIAL (EC 3.6.1.34) (FRAGMENT).
GN   ATP5B.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC       GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC       SUBUNIT.
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   HSC-2DPAGE; P99504; DOG.
DR   INTERPRO; IPR000194; -.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW   ATP synthesis; CF(1); Hydrogen ion transport; Hydrolase; ATP-binding;
KW   Mitochondrion.
FT   UNSURE       8      8
FT   UNSURE      17     19
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1871 MW;  BB9C163FDC60BB42 CRC64;


  Query Match             21.9%;  Score 23;  DB 1;  Length 19;
  Best Local Similarity   41.7%;  Pred. No. 7.1e+02;
  Matches    5;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       8 PRPDGREAVAYR 19
           | | |  |   |
Db       6 PSPKGAAAXXXR 17


RESULT   9
```

```
OS   Halocynthia roretzi (Sea squirt).
OC   Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
OC   Stolidobranchia; Pyuridae; Halocynthia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEMOLYMPH;
RX   MEDLINE; 96321313.
RA   Shishikura F., Abe T., Ohtake S.-I., Tanaka K.;
RT   "Purification and characterization of a 58,000-Da proteinase
RT   inhibitor from the hemolymph of a solitary ascidian, Halocynthia
RT   roretzi.";
RL   Comp. Biochem. Physiol. 114B:1-9(1996).
CC   -!- FUNCTION: STRONGLY INHIBITS TRYPSIN AND PLASMA ENZYME(S) ACTIVITY.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   INTERPRO; IPR000215; -.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Serpin; Serine protease inhibitor; Glycoprotein; Plasma.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1104 MW;  4225C73B1B187AA3 CRC64;


  Query Match             23.8%;  Score 25;  DB 1;  Length 10;
  Best Local Similarity   55.6%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      9 RPDGREAVA 17
          : || | ||
Db      2 KKDGEEKVA 10


RESULT   3
FIBA_BUBAR
ID   FIBA_BUBAR     STANDARD;      PRT;     19 AA.
AC   P14443;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   FIBRINOPEPTIDE A.
OS   Bubalus arnee bubalis (Domestic water buffalo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bubalus.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=ITALIAN BREED;
RX   MEDLINE; 76040091.
RA   Balestrieri C., Colonna G., Irace G.;
RT   "Covalent structure of fibrinopeptides from buffaloes breeding in
RT   Italy.";
RL   Biochim. Biophys. Acta 405:517-521(1975).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   HSSP; P02671; 1BBR.
KW   Blood coagulation; Plasma.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1852 MW;  9BA41F0F55A54CC5 CRC64;


  Query Match             23.8%;  Score 25;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     11 DGREAVAYRF 20
          || :||:  |
Db      2 DGSDAVSGEF 11


RESULT   4
MARI_ALTSP
ID   MARI_ALTSP     STANDARD;      PRT;     14 AA.
AC   P29399;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   MARINOSTATINS C-2, C-1, AND D.
OS   Alteromonas sp. (strain B-10-31).
OC   Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC   Alteromonas.
RN   [1]
RP   SEQUENCE, AND ACTIVE SITE.
RX   MEDLINE; 92176155.
RA   Takano R., Imada C., Kamei K., Hara S.;
RT   "The reactive site of marinostatin, a proteinase inhibitor from
RT   marine Alteromonas sp. B-10-31.";
RL   J. Biochem. 110:856-858(1991).
CC   -!- FUNCTION: INHIBITS SUBTILISIN, CHYMOTRYPSIN, AND ELASTASE, BUT
CC       NOT TRYPSIN.
CC   -!- MISCELLANEOUS: THE SEQUENCE SHOWN IS THAT OF MARINOSTATIN C-2.
KW   Serine protease inhibitor.
FT   MOD_RES      1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   VARIANT      1      2       MISSING (IN MARINOSTATIN C-1).
FT   VARIANT      1      3       MISSING (IN MARINOSTATIN D).
FT   ACT_SITE     6      7       REACTIVE BOND.
SQ   SEQUENCE   14 AA;  1644 MW;  6E7CEEF92EF32E44 CRC64;


  Query Match             22.9%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 TFRFPRPD 11
            | |:|  |
Db      5 TMRYPSDD 12


RESULT   5
FIBA_BUBBU
ID   FIBA_BUBBU     STANDARD;      PRT;     19 AA.
AC   P14442;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   FIBRINOPEPTIDE A.
OS   Bubalus bubalis (Water buffalo).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bubalus.
RN   [1]
RP   SEQUENCE.
RA   Mross G.A., Doolittle R.F.;
RT   "Amino acid sequence studies on artiodacty fibrinopeptides.";
RL   Arch. Biochem. Biophys. 122:674-684(1967).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
DR   HSSP; P02671; 1BBR.
KW   Blood coagulation; Plasma.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1822 MW;  9BA41F1E11A54CC5 CRC64;


  Query Match             22.9%;  Score 24;  DB 1;  Length 19;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:52 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-3
Perfect score: 105
Sequence: 1 VAVTFRFPRPDGREAVAYRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	24.8	15	1	C1QA_RAT	P31720 rattus norv
2	25	23.8	10	1	SPI_HALRO	Q10997 halocynthia
3	25	23.8	19	1	FIBA_BUBAR	P14443 bubalus arn
4	24	22.9	14	1	MARI_ALTSP	P29399 alteromonas
5	24	22.9	19	1	FIBA_BUBBU	P14442 bubalus bub
6	24	22.9	20	1	CAQS_RAT	P19633 rattus norv
7	23	21.9	11	1	CEP1_ACHFU	P22790 achatina fu
8	23	21.9	19	1	ATPB_CANFA	P99504 canis famil
9	23	21.9	19	1	PHSL_DESBN	P13066 desulfovibr
10	22	21.0	10	1	UXA6_CHLTR	P38007 chlamydia t
11	22	21.0	12	1	TKN1_KASMA	P08613 kassina mac
12	22	21.0	19	1	FIBB_HORSE	P14471 equus cabal
13	21	20.0	13	1	LIGA_TRAVE	P20011 trametes ve
14	21	20.0	15	1	MK2A_PALPR	P80409 palomena pr
15	21	20.0	15	1	SODM_STRGR	P80733 streptomyce
16	21	20.0	16	1	MK3_PALPR	P80411 palomena pr
17	21	20.0	18	1	GALS_SALTY	P41030 salmonella
18	20.5	19.5	20	1	MIF_PIG	P80928 sus scrofa
19	20	19.0	9	1	FLA2_TREHY	P80159 treponema h
20	20	19.0	10	1	BPP2_BOTJA	P01022 bothrops ja
21	20	19.0	10	1	PPCK_FASHE	P80525 fasciola he
22	20	19.0	11	1	TKN1_PSEGU	P42986 pseudophryn
23	20	19.0	11	1	TKN2_PSEGU	P42987 pseudophryn
24	20	19.0	11	1	TKN4_PSEGU	P42989 pseudophryn
25	20	19.0	13	1	BPP1_BOTJA	P01020 bothrops ja
26	20	19.0	13	1	ECDE_LYMDI	P80941 lymantria d
27	20	19.0	14	1	LPF2_ECOLI	P06985 escherichia
28	20	19.0	16	1	MK2B_PALPR	P80410 palomena pr
29	20	19.0	17	1	YALA_TRYBB	P17961 trypanosoma
30	20	19.0	20	1	PORC_METTM	P80902 methanobact
31	19	18.1	7	1	UF04_MOUSE	P38642 mus musculu
32	19	18.1	9	1	FAR9_ASCSU	P43172 ascaris suu
33	19	18.1	11	1	BPP3_BOTIN	P30423 bothrops in
34	19	18.1	11	1	BPP4_BOTIN	P30424 bothrops in
35	19	18.1	11	1	BPPB_AGKHA	P01021 agkistrodon
36	19	18.1	11	1	CORZ_PERAM	P11496 periplaneta
37	19	18.1	11	1	TKNA_CHICK	P19850 gallus gall
38	19	18.1	11	1	TKNA_GADMO	P28498 gadus morhu
39	19	18.1	11	1	TKNA_ONCMY	P28499 oncorhynchu
40	19	18.1	11	1	TKNA_RANCA	P22688 rana catesb
41	19	18.1	11	1	TKNA_SCYCA	P41333 scyliorhinu
42	19	18.1	13	1	LIGB_TRAVE	P20012 trametes ve
43	19	18.1	13	1	MLA_ANOCA	P41589 anolis caro
44	19	18.1	13	1	MLA_CAMDR	P01198 camelus dro
45	19	18.1	13	1	YCIA_SALTY	P25944 salmonella

ALIGNMENTS

RESULT 1
C1QA_RAT
ID C1QA_RAT STANDARD; PRT; 15 AA.
AC P31720;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE COMPLEMENT C1Q SUBCOMPONENT, A CHAIN (FRAGMENT).
GN C1QA.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 93218657.
RA Wing M.G., Seilly D.J., Bridgman D.J., Harrison R.A.;
RT "Rapid isolation and biochemical characterization of rat C1 and C1q.";
RL Mol. Immunol. 30:433-440(1993).
CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
CC C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
CC AND S IN THE MOLAR RATION OF 1:2:2. THE C1Q SUBCOMPONENT IS
CC COMPOSED OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED
CC DIMERS OF THE A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-
CC LINKED DIMERS OF THE C CHAIN. IN ADDITION TO THE MAJOR A:B AND C:C
CC DIMER BANDS, RAT, UNLIKE HUMAN C1Q, CONTAINED MINOR DIMER SPECIES.
DR INTERPRO; IPR001073; -.
DR PROSITE; PS01113; C1Q; PARTIAL.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat.
FT DISULFID 4 4 INTERCHAIN (WITH CYS-4 IN B CHAIN).
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1488 MW; 1B3D8000B7793965 CRC64;

Query Match 24.8%; Score 26; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy      10 PDGREAVA 17
             |:|:: ||
Db       7 PNGKDGVA 14
```

RESULT 2
SPI_HALRO
ID SPI_HALRO STANDARD; PRT; 10 AA.
AC Q10997;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SERINE PROTEINASE INHIBITOR (FRAGMENT).

RESULT 13
A34662
Achatina cardio-excitatory peptide-1 - giant African snail
C;Species: Achatina fulica (giant African snail)
C;Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 11-Jul-1997
C;Accession: A34662
R;Fujimoto, K.; Ohta, N.; Yoshida, M.; Kubota, I.; Muneoka, Y.; Kobayashi, M.
Biochem. Biophys. Res. Commun. 167, 777-783, 1990
A;Title: A novel cardio-excitatory peptide isolated from the atria of the African giant
A;Reference number: A34662; MUID:90211261
A;Accession: A34662
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <FUJ>
C;Keywords: amidated carboxyl end
F;11/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 21.9%; Score 23; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 7.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 9 RPDGR 13
|| ||
Db 6 RPQGR 10

RESULT 14
S21241
oligo-1,6-glucosidase (EC 3.2.1.10) - Bacillus "thermoamyloliquefaciens" (fragment)
C;Species: Bacillus "thermoamyloliquefaciens"
C;Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 21-Aug-1998
C;Accession: S21241
R;Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.
Eur. J. Biochem. 205, 249-256, 1992
A;Title: Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-glucosidase I to an
nce and in structural parameters calculated from the amino acid composition.
A;Reference number: S21202; MUID:92209510
A;Accession: S21241
A;Molecule type: protein
A;Residues: 1-15 <SUZ>
A;Experimental source: strain KP1071
C;Superfamily: alpha-glucosidase; alpha-amylase core homology
C;Keywords: glycosidase; hydrolase

Query Match 21.9%; Score 23; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 13 REAVAYR 19
:||| |:
Db 7 KEAVVYQ 13

RESULT 15
S21202
glucan 1,4-alpha-glucosidase (EC 3.2.1.3) - Bacillus stearothermophilus (fragment)
N;Alternate names: exo-alpha-1-4-glucosidase I
C;Species: Bacillus stearothermophilus
C;Date: 22-Nov-1993 #sequence_revision 01-Nov-1996 #text_change 21-Aug-1998
C;Accession: S21202
R;Suzuki, Y.; Yonezawa, K.; Hattori, M.; Takii, Y.
Eur. J. Biochem. 205, 249-256, 1992
A;Title: Assignment of Bacillus thermoamyloliquefaciens KP1071 alpha-glucosidase I to an
nce and in structural parameters calculated from the amino acid composition.
A;Reference number: S21202; MUID:92209510
A;Accession: S21202
A;Molecule type: protein
A;Residues: 1-15 <SUZ>
A;Experimental source: ATCC 12016
C;Superfamily: alpha-glucosidase; alpha-amylase core homology
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 21.9%; Score 23; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 13 REAVAYR 19
:| |||:
Db 7 KEGVAYQ 13

Search completed: December 21, 2000, 08:30:03
Job time: 267 sec

R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensiona
A;Reference number: PA0001
A;Accession: PA0003
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: leaf and callus
C;Superfamily: nucleoside-diphosphate kinase
C;Keywords: phosphotransferase

Query Match 22.9%; Score 24; DB 2; Length 15;
Best Local Similarity 55.6%; Pred. No. 7.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 TFRFPRPDG 12
|| :|||
Db 6 TFIAIKPDG 14

RESULT 9
A28965
ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain - spinach (fragments)
C;Species: Spinacia oleracea (spinach)
C;Date: 22-Dec-1988 #sequence_revision 22-Dec-1988 #text_change 23-Feb-1997
C;Accession: A28965
R;Mulligan, R.M.; Houtz, R.L.; Tolbert, N.E.
Proc. Natl. Acad. Sci. U.S.A. 85, 1513-1517, 1988
A;Title: Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/oxy
cetylated proline.
A;Reference number: A28965; MUID:88144466
A;Accession: A28965
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <MUL>
C;Keywords: carbon-carbon lyase; carboxy-lyase; chloroplast

Query Match 22.9%; Score 24; DB 2; Length 15;
Best Local Similarity 45.5%; Pred. No. 7.4e+02;
Matches 5; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 VAVTFRFPRPD 11
| | || |
Db 3 VGFKFEFPAMD 13

RESULT 10
S09722
2S albumin small chain 1 nIV - rape (fragments)
C;Species: Brassica napus (rape)
C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998
C;Accession: S09722
R;Monsalve, R.I.; Menendez-Arias, L.; Lopez-Otin, C.; Rodriguez, R.
FEBS Lett. 263, 209-212, 1990
A;Title: beta-Turns as structural motifs for the proteolytic processing of seed proteins
A;Reference number: S09720; MUID:90242974
A;Accession: S09722
A;Molecule type: protein
A;Residues: 1-9;10-18 <MON>
A;Experimental source: seed

Query Match 22.9%; Score 24; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 FRFPRPDG 12
|| |: |
Db 5 FRIPKQPG 12

RESULT 11
A31049
calsequestrin, fast skeletal muscle - rat (fragment)
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; lamin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 16-Jul-1999
C;Accession: A31049; S46193; S46194; S46195
R;Hall, D.E.; Frazer, K.A.; Hann, B.C.; Reichardt, L.F.
J. Cell Biol. 107, 687-697, 1988
A;Title: Isolation and characterization of a laminin-binding protein from rat and chi
A;Reference number: A92751; MUID:88331073
A;Accession: A31049
A;Molecule type: protein
A;Residues: 1-20 <HAL>
R;Volpe, P.; Martini, A.; Furlan, S.; Meldolesi, J.
Biochem. J. 301, 465-469, 1994
A;Title: Calsequestrin is a component of smooth muscles: the skeletal- and cardiac-mu
A;Reference number: S46193; MUID:94318050
A;Accession: S46193
A;Molecule type: protein
A;Residues: 1-9 <VOL>
A;Accession: S46194
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'XXX',4-10,'X',12 <VO2>
A;Accession: S46195
A;Status: preliminary
A;Molecule type: protein
A;Residues: 'XE',3-12 <VO3>
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding pro
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound t
C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in
C;Superfamily: calsequestrin
C;Keywords: calcium binding; glycoprotein; skeletal muscle

Query Match 22.9%; Score 24; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 7 FPRPDGREAV 16
|| || : |
Db 6 FPEYDGVDRV 15

RESULT 12
A58839
ribulose-bisphosphate carboxylase activase I - rice (fragment)
N;Alternate names: RUBISCO activase I, 44K
C;Species: Oryza sativa (rice)
C;Date: 17-Jun-1998 #sequence_revision 02-Jul-1998 #text_change 02-Jul-1998
C;Accession: A58839
R;To, K.Y.; Suen, D.F.; Chen, L.F.O.; Chen, S.C.G.
submitted to the Protein Sequence Database, June 1998
A;Description: Characterization of rice leaf cDNA encoding rubisco activase.
A;Reference number: A58839
A;Accession: A58839
A;Molecule type: protein
A;Residues: 1-15 <TOK>
A;Experimental source: strain Tainong 67; leaf
A;Note: RUBISCO activase I has an additional 33 residues at the carboxyl end compared
C;Keywords: leaf

Query Match 22.4%; Score 23.5; DB 2; Length 15;
Best Local Similarity 53.8%; Pred. No. 8.9e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

Qy 5 FRFPRPDG-REAV 16
|: | || |||
Db 1 FQAPTGDGTHEAV 13

```
          || || ||:
Db      8 FPVVDGCEAM 17


RESULT   3
PT0091
H+-transporting ATP synthase (EC 3.6.1.34) alpha chain - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 21-Aug-1998 #sequence_revision 21-Aug-1998 #text_change 21-Aug-1998
C;Accession: PT0091
R;Kawakami, T.; Uchida, T.; Sakai, T.; Kamo, M.; Morimasa, T.; Tsugita, A.
submitted to JIPID, July 1998
A;Description: Proteome analysis of mouse brain.
A;Reference number: PT0091
A;Accession: PT0091
A;Molecule type: protein
A;Residues: 1-15 <KAW>
A;Experimental source: brain, striatum
C;Keywords: hydrolase


  Query Match             24.8%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PDGREA 15
           | ||||
Db       3 PPGREA 8


RESULT   4
E39419
collagen alpha 5(IV) chain - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 19-Oct-1995
C;Accession: E39419
R;Gunwar, S.; Ballester, F.; Kalluri, R.; Timoneda, J.; Chonko, A.M.; Edwards, S.J.; Noe
J. Biol. Chem. 266, 15318-15324, 1991
A;Title: Glomerular basement membrane. Identification of dimeric subunits of the noncoll
A;Reference number: A39419; MUID:91332055
A;Accession: E39419
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <GUN>
C;Superfamily: collagen alpha 1(IV) chain
C;Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix


  Query Match             24.8%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   45.5%;  Pred. No. 4.6e+02;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      10 PDGREAVAYRF 20
           | |  :||: |
Db       2 PPGTSSVAHGF 12


RESULT   5
S09721
2S albumin small chain nIII - rape (fragments)
C;Species: Brassica napus (rape)
C;Date: 19-Mar-1997 #sequence_revision 13-Mar-1998 #text_change 13-Mar-1998
C;Accession: S09721
R;Monsalve, R.I.; Menendez-Arias, L.; Lopez-Otin, C.; Rodriguez, R.
FEBS Lett. 263, 209-212, 1990
A;Title: beta-Turns as structural motifs for the proteolytic processing of seed proteins
A;Reference number: S09720; MUID:90242974
A;Accession: S09721
A;Molecule type: protein
A;Residues: 1-9;10-14 <MON>
A;Experimental source: seed


  Query Match             23.8%;  Score 25;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 4.6e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 FRFPRPDG 12
            || |:  |
Db       5 FRIPKQSG 12


RESULT   6
I54264
rhodopsin single base deletion frame shift mutant - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 08-Nov-1996 #text_change 20-Apr-2000
C;Accession: I54264
R;Horn, M.; Humphries, P.; Kunisch, M.; Marchese, C.; Apfelstedt-Sylla, E.; Fugi, L.;
Hum. Genet. 90, 255-257, 1992
A;Title: Deletions in exon 5 of the human rhodopsin gene causing a shift in the readi
A;Reference number: I54264; MUID:93138610
A;Accession: I54264
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-19 <HOR>
A;Cross-references: GB:S55843; NID:g266287; PIDN:AAB25673.1; PID:g266288
C;Genetics:
A;Gene: GDB:RHO
A;Cross-references: GDB:120347
A;Map position: 3q21.3-3q24


  Query Match             23.8%;  Score 25;  DB 4;  Length 19;
  Best Local Similarity   80.0%;  Pred. No. 6.4e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 RFPRP 10
            |:|||
Db       5 RWPRP 9


RESULT   7
S68402
NAD(+)-glycohydrolase - bovine (fragments)
C;Species: Bos primigenius taurus (cattle)
C;Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 12-Dec-1997
C;Accession: S68402
R;Zhang, J.; Ziegler, M.; Schneider, R.; Klocker, H.; Auer, B.; Schweiger, M.
FEBS Lett. 377, 530-534, 1995
A;Title: Identification and purification of a bovine liver mitochondrial NAD(+)-glyco
A;Reference number: S68402; MUID:96140583
A;Accession: S68402
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7;8-12 <ZHA>
C;Genetics:
A;Genome: nuclear
C;Keywords: mitochondrion; NAD


  Query Match             22.9%;  Score 24;  DB 2;  Length 12;
  Best Local Similarity   50.0%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 DGREAVAY 18
           |||: :|:
Db       5 DGRKELAH 12


RESULT   8
PA0003
nucleoside-diphosphate kinase (EC 2.7.4.6) - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 18-Nov-1994 #sequence_revision 06-Jan-1995 #text_change 15-Mar-1996
C;Accession: PA0003
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:01 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-3
Perfect score: 105
Sequence: 1 VAVTFRFPRPDGREAVAYRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	25.7	13	2	I84603	deoxynucleotidyltr
2	27	25.7	17	2	C37520	glutathione transf
3	26	24.8	15	2	PT0091	H+-transporting AT
4	26	24.8	20	2	E39419	collagen alpha 5(I
5	25	23.8	14	2	S09721	2S albumin small c
6	25	23.8	19	4	I54264	rhodopsin single b
7	24	22.9	12	2	S68402	NAD(+)-glycohydrol
8	24	22.9	15	2	PA0003	nucleoside-diphosp
9	24	22.9	15	2	A28965	ribulose-bisphosph
10	24	22.9	18	2	S09722	2S albumin small c
11	24	22.9	20	2	A31049	calsequestrin, fas
12	23.5	22.4	15	2	A58839	ribulose-bisphosph
13	23	21.9	11	2	A34662	Achatina cardio-ex
14	23	21.9	15	2	S21241	oligo-1,6-glucosid
15	23	21.9	15	2	S21202	glucan 1,4-alpha-g
16	23	21.9	17	1	A61339	vespulakinin 1 - e
17	23	21.9	19	2	H27480	hydrogenase (EC 1.
18	23	21.9	20	2	S13274	ferredoxin [2Fe-2S
19	22	21.0	12	2	S10059	tachykinin - Afric
20	22	21.0	13	2	T08533	hypothetical prote
21	22	21.0	13	2	S22995	hypothetical prote
22	22	21.0	16	2	PS0210	28K protein 4209 -
23	22	21.0	17	2	B25348	UDPglucose--glycog
24	22	21.0	19	2	A92058	pepsin-like protei
25	21	20.0	9	2	S66607	quinoline 2-oxidor
26	21	20.0	10	2	A61007	hementin (EC 3.4.-
27	21	20.0	10	2	PH0926	T-cell receptor be
28	21	20.0	11	2	H54346	pyruvate synthase
29	21	20.0	13	2	S04013	lignin peroxidase
30	21	20.0	14	2	S60353	amylopullulanase -
31	21	20.0	14	2	I54284	C1-inhibitor - hum
32	21	20.0	15	2	F57789	gallbladder stone
33	21	20.0	17	2	A34835	ribosomal protein
34	21	20.0	19	2	PQ0548	capsid protein VP2
35	20	19.0	8	2	S66646	cardioacceleratory
36	20	19.0	9	2	A37027	macrophage chemota
37	20	19.0	9	2	I46023	growth hormone rec
38	20	19.0	10	1	XAVI6B	angiotensin-conver
39	20	19.0	11	2	B60409	kassinin-like pept
40	20	19.0	11	2	C60409	kassinin-like pept
41	20	19.0	11	2	E60409	substance P-like p
42	20	19.0	11	2	D56979	collagen alpha 1(I
43	20	19.0	13	1	XAVI9B	angiotensin-conver
44	20	19.0	13	2	A37196	bradykinin-potenti
45	20	19.0	14	1	LFECFS	pheST operon leade

ALIGNMENTS

RESULT 1
I84603
deoxynucleotidyltransferase - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 29-Aug-1997
C;Accession: I84603
R;Koiwai, O.; Kaneda, T.; Morishita, R.
Biochem. Biophys. Res. Commun. 144, 185-190, 1987
A;Title: Analysis of human terminal deoxynucleotidyl transferase cDNA expressible in
A;Reference number: I45884; MUID:87213162
A;Accession: I84603
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-13 <RES>
A;Cross-references: GB:M26144; NID:g951194; PID:g951195
C;Genetics:
A;Gene: GDB:DNTT
A;Cross-references: GDB:119100; OMIM:187410
A;Map position: 10q23-10q24

Query Match 25.7%; Score 27; DB 2; Length 13;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      9 RPDGRE 14
          :||||:
Db      3 QPDGRQ 8
```

RESULT 2
C37520
glutathione transferase (EC 2.5.1.18) MII - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 30-Sep-1993
C;Accession: C37520; N24735
R;Mannervik, B.; Alin, P.; Guthenberg, C.; Jensson, H.; Tahir, M.K.; Warholm, M.; Jor
Proc. Natl. Acad. Sci. U.S.A. 82, 7202-7206, 1985
A;Title: Identification of three classes of cytosolic glutathione transferase common
A;Reference number: A24735; MUID:86042634
A;Accession: C37520
A;Molecule type: protein
A;Residues: 1-17 <MAN>
C;Superfamily: glutathione transferase
C;Keywords: transferase

Query Match 25.7%; Score 27; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 2.6e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy      7 FPRPDGREAV 16
```

```
Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT  14
5196404-4
;Patent No. 5196404
;     APPLICANT: MARAGANORE, JOHN M.; FENTON II, JOHN M.; KLINE, TONI
;     TITLE OF INVENTION: INHIBITORS OF THROMBIN
;     NUMBER OF SEQUENCES: 18
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/549,388
;       FILING DATE: 06-JUL-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 395,482
;       FILING DATE: 18-AUG-1989
;SEQ ID NO:4:
;       LENGTH: 20
5196404-4


  Query Match             28.6%;  Score 30;  DB 5;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT  15
5425936-1
;Patent No. 5425936
;     APPLICANT: MARAGANORE, JOHN M.;JABLONSKI, JO-ANN M.;BOURDON,
;PAUL R.
;     TITLE OF INVENTION: INHIBITORS OF THROMBIN
;     NUMBER OF SEQUENCES: 8
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/924,549
;       FILING DATE: 31-JUL-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 652,929
;       FILING DATE: 08-FEB-1991
;       APPLICATION NUMBER: 549,388
;       FILING DATE:06-JUL-1990
;       APPLICATION NUMBER: 395,482
;       FILING DATE: 18-AUG-1989
;SEQ ID NO:1:
;       LENGTH: 20
5425936-1


  Query Match             28.6%;  Score 30;  DB 5;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


Search completed: December 21, 2000, 08:31:45
Job time: 368 sec
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,052
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5788960nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  90,1104-V
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312 715 1000
;      TELEFAX:  312 715 1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Phe
;      OTHER INFORMATION:  /note= "The amino terminal phenylalanine residue
;      OTHER INFORMATION:  is in the D stereochemical configuration"
US-08-463-052-29


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT  12
US-08-480-551-29
; Sequence 29, Application US/08480551
; Patent No. 5811394
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T
;    TITLE OF INVENTION:  Technetium-99m Labeled Polypeptides for
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,551
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/264,176
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/653,012
;      FILING DATE:  08-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  90,1104
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312 715 1000
;      TELEFAX:  312 715 1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Phe
;      OTHER INFORMATION:  /note= "The amino terminal phenylalanine residue
;      OTHER INFORMATION:  is in the D stereochemical configuration"
US-08-480-551-29


  Query Match             28.6%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT  13
PCT-US91-09108-18
; Sequence 18, Application PC/TUS9109108
;  GENERAL INFORMATION:
;    APPLICANT:  Biogen, Inc.
;    TITLE OF INVENTION:  *\ ' BIFUNCTIONAL INHIBITORS OF THROMBIN ANDPLATELET ACTIVA
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Neave
;      STREET:  875 Third Avenue - 29th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/09108
;      FILING DATE:  19911205
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haley Jr., James F.
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  B154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 715-0600
;      TELEFAX:  (212) 715-0674
;      TELEX:  14-8367
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-09108-18


  Query Match             28.6%;  Score 30;  DB 4;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  AMINO ACID
;      STRANDEDNESS:
;      TOPOLOGY:  Linear
US-08-058-699-9


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      7 FPRPDG 12
            |||| |
Db      1 FPRPGG 6


RESULT   9
US-08-464-456-29
; Sequence 29, Application US/08464456
; Patent No. 5681541
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T
;    TITLE OF INVENTION:  Technetium-99m Labeled Polypeptides for
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10 South Wacker Drive Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/464,456
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5681541nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  90,1104-V
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312 715 1000
;      TELEFAX:  312 715 1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  1
;      OTHER INFORMATION:  /label= D-Phe
;      OTHER INFORMATION:  /note= "The amino terminal phenylalanine residue
;      OTHER INFORMATION:  is in the D stereochemical configuration"
US-08-464-456-29


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      7 FPRPDG 12
            |||| |
Db      1 FPRPGG 6


RESULT  10
US-08-286-748B-6
; Sequence 6, Application US/08286748B
; Patent No. 5759542
;  GENERAL INFORMATION:
;    APPLICANT:  Victor Gurewich
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE DELIVERY
;    TITLE OF INVENTION:  OF DRUGS BY PLATELETS FOR THE TREATMENT OF
;    TITLE OF INVENTION:  CARDIOVASCULAR AND OTHER DISEASES
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/286,748B
;      FILING DATE:  August 5, 1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  J. Peter Fasse
;      REGISTRATION NUMBER:  32,983
;      REFERENCE/DOCKET NUMBER:  04547/013001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-286-748B-6


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      7 FPRPDG 12
            |||| |
Db      1 FPRPGG 6


RESULT  11
US-08-463-052-29
; Sequence 29, Application US/08463052
; Patent No. 5788960
;  GENERAL INFORMATION:
;    APPLICANT:  Dean, Richard T
;    TITLE OF INVENTION:  Technetium-99m Labeled Polypeptides for
;    TITLE OF INVENTION:  Imaging
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10 South Wacker Drive Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
```

```
; Patent No. 5242810
;  GENERAL INFORMATION:
;    APPLICANT:  Maraganore, John M.
;    APPLICANT:  Chao, Betty H.
;    APPLICANT:  Strauch, Kathryn L.
;    APPLICANT:  Thompson, Jeffrey S.
;    TITLE OF INVENTION:  BIFUNCTIONAL INHIBITORS OF THROMBIN AND
;    TITLE OF INVENTION:  PLATELET ACTIVATION
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Neave
;      STREET:  875 Third Avenue - 29th Floor
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.24
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/623,611
;      FILING DATE:  19901207
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haley Jr., James F.
;      REGISTRATION NUMBER:  27,794
;      REFERENCE/DOCKET NUMBER:  B154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 715-0600
;      TELEFAX:  (212) 715-0674
;      TELEX:  14-8367
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-07-623-611-18


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT   7
US-07-831-780-2
; Sequence 2, Application US/07831780
; Patent No. 5371184
;  GENERAL INFORMATION:
;    APPLICANT:  Dunn, Jeffrey
;    APPLICANT:  Lyle, Leon R
;    APPLICANT:  Rajagopalan, Raghavan
;    TITLE OF INVENTION:  Radiolabelled Peptide Compounds
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rita D. Vacca
;      STREET:  Mallinckrodt Medical, Inc., 675 McDonnell
;      STREET:  Blvd.
;      CITY:  St. Louis
;      STATE:  MO
;      COUNTRY:  USA
;      ZIP:  63134
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/831,780
;      FILING DATE:  19920502
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vacca, Rita D.
;      REGISTRATION NUMBER:  33,624
;      REFERENCE/DOCKET NUMBER:  0754
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  314-895-7215
;      TELEFAX:  314-895-2156
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Hirudin
;      STRAIN:  Human
US-07-831-780-2


  Query Match             28.6%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 72;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT   8
US-08-058-699-9
; Sequence 9, Application US/08058699
; Patent No. 5443827
;  GENERAL INFORMATION:
;    APPLICANT:  Edgar Haber
;    APPLICANT:  Christoph Bode
;    APPLICANT:  Marschall S. Runge
;    TITLE OF INVENTION:  FIBRIN-TARGETED INHIBITORS OF
;    TITLE OF INVENTION:  THROMBIN
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/058,699
;      FILING DATE:  19930503
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  N/A
;      FILING DATE:  N/A
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Janis K. Fraser, Ph.D.
;      REGISTRATION NUMBER:  34,819
;      REFERENCE/DOCKET NUMBER:  05433/004001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  9:
```

```
RESULT   2
5433940-7
;Patent No. 5433940
;     APPLICANT: MARAGANORE, JOHN M.; FENTON II, JOHN M.; KLINE, TONI
;     TITLE OF INVENTION: INHIBITORS OF THROMBIN
;     NUMBER OF SEQUENCES: 25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/834,259
;       FILING DATE: 17-AUG-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 548,388
;       FILING DATE: 06-JUL-1989
;       APPLICATION NUMBER: 395,482
;       FILING DATE: 18-AUG-1989
;SEQ ID NO:7:
;       LENGTH: 10
5433940-7


  Query Match             28.6%;  Score 30;  DB 5;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 33;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT   3
US-08-101-041A-4
; Sequence 4, Application US/08101041A
; Patent No. 5541101
;  GENERAL INFORMATION:
;     APPLICANT:  SAJI, Fumitaka
;     APPLICANT:  AZUMA, Chihiro
;     APPLICANT:  KIMURA, Tadashi
;     TITLE OF INVENTION:  ANTI-OXYTOCIN RECEPTOR ANTIBODIES AND
;     TITLE OF INVENTION:  METHODS FOR THEIR PRODUCTION
;     NUMBER OF SEQUENCES:  7
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Burns, Doane, Swecker & Mathis
;       STREET:  P.O. Box 1404
;       CITY:  Alexandria
;       STATE:  Virginia
;       COUNTRY:  United States
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/101,041A
;       FILING DATE:  03-AUG-1993
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  JP 4-206854
;       FILING DATE:  03-AUG-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Teskin, Robin L
;       REGISTRATION NUMBER:  35,030
;       REFERENCE/DOCKET NUMBER:  002258-008
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 836-6620
;       TELEFAX:  (703) 836-2021
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  1..18
;       OTHER INFORMATION:  /note= "Amino acids 102-119 of the
;       OTHER INFORMATION:  oxytocin receptor polypeptide."
US-08-101-041A-4


  Query Match             28.6%;  Score 30;  DB 1;  Length 18;
  Best Local Similarity   75.0%;  Pred. No. 64;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TFRFPRPD 11
           ||||  ||
Db       1 TFRFYGPD 8


RESULT   4
5196404-6
;Patent No. 5196404
;     APPLICANT: MARAGANORE, JOHN M.; FENTON II, JOHN M.; KLINE, TONI
;     TITLE OF INVENTION: INHIBITORS OF THROMBIN
;     NUMBER OF SEQUENCES: 18
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/549,388
;       FILING DATE: 06-JUL-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 395,482
;       FILING DATE: 18-AUG-1989
;SEQ ID NO:6:
;       LENGTH: 18
5196404-6


  Query Match             28.6%;  Score 30;  DB 5;  Length 18;
  Best Local Similarity   83.3%;  Pred. No. 64;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT   5
5433940-2
;Patent No. 5433940
;     APPLICANT: MARAGANORE, JOHN M.; FENTON II, JOHN M.; KLINE, TONI
;     TITLE OF INVENTION: INHIBITORS OF THROMBIN
;     NUMBER OF SEQUENCES: 25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/834,259
;       FILING DATE: 17-AUG-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 548,388
;       FILING DATE: 06-JUL-1989
;       APPLICATION NUMBER: 395,482
;       FILING DATE: 18-AUG-1989
;SEQ ID NO:2:
;       LENGTH: 19
5433940-2


  Query Match             28.6%;  Score 30;  DB 5;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 68;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0,  Gaps    0;

Qy       7 FPRPDG 12
           |||| |
Db       1 FPRPGG 6


RESULT   6
US-07-623-611-18
; Sequence 18, Application US/07623611
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:43 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-3
Perfect score: 105
Sequence: 1 VAVTFRFPRPDGREAVAYRF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	28.6	7	2	US-08-713-885-6	Sequence 6, Appli
2	30	28.6	10	5	5433940-7	Patent No. 5433940
3	30	28.6	18	1	US-08-101-041A-4	Sequence 4, Appli
4	30	28.6	18	5	5196404-6	Patent No. 5196404
5	30	28.6	19	5	5433940-2	Patent No. 5433940
6	30	28.6	20	1	US-07-623-611-18	Sequence 18, Appl
7	30	28.6	20	1	US-07-831-780-2	Sequence 2, Appli
8	30	28.6	20	1	US-08-058-699-9	Sequence 9, Appli
9	30	28.6	20	1	US-08-464-456-29	Sequence 29, Appl
10	30	28.6	20	1	US-08-286-748B-6	Sequence 6, Appli
11	30	28.6	20	1	US-08-463-052-29	Sequence 29, Appl
12	30	28.6	20	2	US-08-480-551-29	Sequence 29, Appl
13	30	28.6	20	4	PCT-US91-09108-18	Sequence 18, Appl
14	30	28.6	20	5	5196404-4	Patent No. 5196404
15	30	28.6	20	5	5425936-1	Patent No. 5425936
16	29	27.6	10	3	US-08-159-339A-102	Sequence 102, App
17	29	27.6	11	3	US-08-893-526A-17	Sequence 17, Appl
18	29	27.6	15	3	US-08-630-916A-96	Sequence 96, Appl
19	28.5	27.1	20	1	US-08-173-116-2	Sequence 2, Appli
20	28	26.7	13	3	US-08-974-775-12	Sequence 12, Appl
21	28	26.7	14	3	US-08-974-775-11	Sequence 11, Appl
22	27	25.7	10	1	US-07-602-847C-1	Sequence 1, Appli
23	27	25.7	14	2	US-08-622-720A-10	Sequence 10, Appl
24	27	25.7	14	2	US-08-622-720A-21	Sequence 21, Appl
25	27	25.7	14	3	US-08-974-775-39	Sequence 39, Appl
26	27	25.7	15	1	US-07-602-847C-6	Sequence 6, Appli
27	27	25.7	15	2	US-08-683-262B-70	Sequence 70, Appl
28	27	25.7	17	5	5219991-5	Patent No. 5219991
29	26	24.8	12	2	US-08-556-597-138	Sequence 138, App
30	26	24.8	13	5	5196523-19	Patent No. 5196523
31	26	24.8	14	2	US-08-531-662B-21	Sequence 21, Appl
32	26	24.8	14	3	US-08-669-161A-21	Sequence 21, Appl
33	26	24.8	14	3	US-08-669-161A-24	Sequence 24, Appl
34	26	24.8	15	1	US-07-978-895-8	Sequence 8, Appli
35	26	24.8	15	2	US-08-473-119-8	Sequence 8, Appli
36	26	24.8	15	2	US-08-475-352-8	Sequence 8, Appli
37	26	24.8	15	2	US-08-726-306A-176	Sequence 176, App
38	26	24.8	16	2	US-08-448-600-5	Sequence 5, Appli
39	26	24.8	16	3	US-08-937-610-9	Sequence 9, Appli
40	26	24.8	19	1	US-08-392-828C-18	Sequence 18, Appl
41	26	24.8	19	3	US-09-330-945-18	Sequence 18, Appl
42	26	24.8	20	1	US-07-956-848A-16	Sequence 16, Appl
43	26	24.8	20	1	US-08-471-956-16	Sequence 16, Appl
44	26	24.8	20	2	US-08-637-759B-413	Sequence 413, App
45	26	24.8	20	3	US-08-871-355A-413	Sequence 413, App

ALIGNMENTS

```
RESULT 1
US-08-713-885-6
; Sequence 6, Application US/08713885
; Patent No. 5985833
; GENERAL INFORMATION:
;   APPLICANT: Mosesson, Michael W.
;   APPLICANT: Meh, David A.
;   TITLE OF INVENTION: THROMBIN INHIBITOR
;   NUMBER OF SEQUENCES: 25
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Quarles & Brady
;     STREET: 411 East Wisconsin Avenue
;     CITY: Milwaukee
;     STATE: Wisconsin
;     COUNTRY: U.S.A.
;     ZIP: 53202-4497
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/713,885
;     FILING DATE:
;     CLASSIFICATION: 514
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Baker, Jean C.
;     REGISTRATION NUMBER: 35,433
;     REFERENCE/DOCKET NUMBER: 960296.93740
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (414) 277-5709
;     TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 6:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-713-885-6


  Query Match              28.6%;  Score 30;  DB 2;  Length 7;
  Best Local Similarity    100.0%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 RFPRP 10
           |||||
Db       3 RFPRP 7
```

```
FT   NON_TER       1      1
SQ   SEQUENCE   22 AA;  2292 MW;  2E05FE169844236D CRC64;


  Query Match             19.4%;  Score 26;  DB 6;  Length 22;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      9 MDFGFPE 15
          || | ||
Db      3 MDIGGPE 9


RESULT  12
Q46081
ID   Q46081      PRELIMINARY;      PRT;    24 AA.
AC   Q46081;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NITROGENASE 3 SUBUNIT H (FRAGMENT).
GN   ANFH.
OS   Clostridium hungatei.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B3B;
RA   Chen T., Leschine S.B.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U59415; AAB02934.1; -.
DR   INTERPRO; IPR000392; -.
DR   PFAM; PF00142; fer4_NifH; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   24 AA;  2752 MW;  C014801651E11BFC CRC64;


  Query Match             19.4%;  Score 26;  DB 2;  Length 24;
  Best Local Similarity   31.2%;  Pred. No. 2e+03;
  Matches    5;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     10 DFGFPEHLLVDFLQSL 25
          :|  |: | :| |: :
Db      1 EFVIPKPLTMDQLEDM 16


RESULT  13
P82401
ID   P82401      PRELIMINARY;      PRT;    25 AA.
AC   P82401;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AUREIN 5.1.
OS   Litoria raniformis, and Litoria aurea (Australian frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: HAS NO ANTIMICROBIAL OR ANTICANCER ACTIVITY.
SQ   SEQUENCE   25 AA;  2547 MW;  15C6169CD98AFC27 CRC64;


  Query Match             19.4%;  Score 26;  DB 13;  Length 25;
  Best Local Similarity   38.9%;  Pred. No. 2.1e+03;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPEHLLVDFLQ 23
          || :  |   :|:|| |:
Db      2 LLDIVTGLLGNLIVDVLK 19


RESULT  14
O94554
ID   O94554      PRELIMINARY;      PRT;    16 AA.
AC   O94554;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CALMODULIN KINASE 2 (FRAGMENT).
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alemany V., Algue R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57982; AAD09466.1; -.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1846 MW;  4A673B1F4C3288D9 CRC64;


  Query Match             18.7%;  Score 25;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     11 FGFP 14
          ||||
Db      5 FGFP 8


RESULT  15
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;


  Query Match             18.7%;  Score 25;  DB 13;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+03;
  Matches    6;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPEHLLVD 20
          ||    | |  | :|
Db      1 LLAQGLGXPTALALD 15


Search completed: December 21, 2000, 08:40:02
Job time: 125 sec
```

```
AC   Q9TRQ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN PEPTIDE L-8, CAP-50=ANNEXIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92317074.
RA   Mizutani A., Usuda N., Tokumitsu H., Minami H., Yasui K.,
RA   Kobayashi R., Hidaka H.;
RT   "CAP-50, a newly identified annexin, localizes in nuclei of cultured
RT   fibroblast 3Y1 cells.";
RL   J. Biol. Chem. 267:13498-13504(1992).
SQ   SEQUENCE   14 AA;  1446 MW;  C8322EB96DD9C6C6 CRC64;


  Query Match             20.1%;  Score 27;  DB 6;  Length 14;
  Best Local Similarity   41.7%;  Pred. No. 7.7e+02;
  Matches    5;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      11 FGFPEHLLVDFL 22
           ||  |  ::| |
Db       2 FGTDEQAIIDXL 13


RESULT   8
Q9S8E1
ID   Q9S8E1      PRELIMINARY;      PRT;    22 AA.
AC   Q9S8E1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ANNEXIN (FRAGMENT).
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Capsicum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95353219.
RA   Hoshino T., Mizutani A., Chida M., Hidaka H., Mizutani J.;
RT   "Plant annexin form homodimer during Ca(2+)-dependent liposome
RT   aggregation.";
RL   Biochem. Mol. Biol. Int. 35:749-755(1995).
SQ   SEQUENCE   22 AA;  2465 MW;  A36D11A1FE311F0F CRC64;


  Query Match             20.1%;  Score 27;  DB 10;  Length 22;
  Best Local Similarity   41.7%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 FGFPEHLLVDFL 22
           :|  | |::| |
Db       3 WGTDEKLIIDIL 14


RESULT   9
Q9TRC2
ID   Q9TRC2      PRELIMINARY;      PRT;    20 AA.
AC   Q9TRC2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   HEMOPROTEIN P-30 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94245717.
RA   Escriou V., Laporte F., Garin J., Brandolin G., Vignais P.V.;
RT   "Purification and physical properties of a novel type of cytochrome b
RT   from rabbit peritoneal neutrophils.";
RL   J. Biol. Chem. 269:14007-14014(1994).
SQ   SEQUENCE   20 AA;  2198 MW;  2301B618163DA419 CRC64;


  Query Match             19.8%;  Score 26.5;  DB 6;  Length 20;
  Best Local Similarity   45.0%;  Pred. No. 1.4e+03;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    5;  Gaps    1;

Qy       3 GFLLLQMDFGFPEHLLVDFL 22
           |||:| :       ||| ||
Db       6 GFLVLLV-----SALLVGFL 20


RESULT  10
Q9ZG55
ID   Q9ZG55      PRELIMINARY;      PRT;    21 AA.
AC   Q9ZG55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ATP-BINDING PROTEIN (FRAGMENT).
GN   RECF.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087306; AAD04082.1; -.
KW   ATP-binding.
FT   NON_TER       1      1
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2336 MW;  0185D9AC428276D9 CRC64;


  Query Match             19.8%;  Score 26.5;  DB 2;  Length 21;
  Best Local Similarity   41.7%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       6 LLQMDFGFPEHL 17
           :||:   ||:|:
Db       8 ILQL-ISFPKHI 18


RESULT  11
O02830
ID   O02830      PRELIMINARY;      PRT;    22 AA.
AC   O02830;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PRO ALPHA1 TYPE II COLLAGEN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96377339.
RA   Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA   Vuorio E.;
RT   "Evidence for insufficient chondrocytic differentiation during repair
RT   of full-thickness defects of articular cartilage.";
RL   Matrix Biol. 15:39-47(1996).
DR   EMBL; S83370; AAB50773.1; -.
DR   INTERPRO; IPR000885; -.
DR   PFAM; PF01410; COLFI; 1.
```

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92250478.
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
SQ   SEQUENCE   19 AA;  2018 MW;  9A54062504B8322E CRC64;


  Query Match             23.1%;  Score 31;  DB 6;  Length 19;
  Best Local Similarity   43.8%;  Pred. No. 2.7e+02;
  Matches    7;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     11 FGFPEHLLVDFLQSLS 26
          ||  |  ::| | | |
Db      2 FGTDEQAIIDXLGSRS 17


RESULT   3
Q9UJH1
ID   Q9UJH1      PRELIMINARY;      PRT;    17 AA.
AC   Q9UJH1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ436M11.2 (RETINOSCHISIS (X-LINKED, JUVENILE) 1 (XLRS1)) (FRAGMENT).
GN   RS1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Grafham D.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z94056; CAB40073.1; -.
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  2029 MW;  8100E1B78C52C7FB CRC64;


  Query Match             22.8%;  Score 30.5;  DB 4;  Length 17;
  Best Local Similarity   58.3%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      2 DGFLLLQMDFGF 13
          :|||||  : ||:
Db      6 EGFLLLLL-FGY 16


RESULT   4
P70861
ID   P70861      PRELIMINARY;      PRT;    21 AA.
AC   P70861;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   THDF (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE; 97312006.
RA   Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT   "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT   sigma 70-like promoter.";
RL   Microbiology 143:1681-1690(1997).
DR   EMBL; U62901; AAB62742.1; -.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2432 MW;  F33E1EC548BD5B33 CRC64;


  Query Match             22.4%;  Score 30;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+02;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      9 MDFGFPEHLL 18
          ::| ||| :|
Db      7 IEFDFPEGIL 16


RESULT   5
Q9TWH5
ID   Q9TWH5      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   40 KDA GAP JUNCTION PROTEIN (FRAGMENT).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Heliothis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95347000.
RA   Ryerse J.S.;
RT   "Immunocytochemical, electrophoresis, and immunoblot analysis of
RT   Heliothis virescens gap junctions isolated in the presence and absence
RT   of protease inhibitors.";
RL   Cell Tissue Res. 281:179-186(1995).
SQ   SEQUENCE   20 AA;  2304 MW;  A298D3EB3E89586B CRC64;


  Query Match             21.6%;  Score 29;  DB 5;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 5.7e+02;
  Matches    6;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGFPEHLLVDF 21
          ::  :| |: | |  ||
Db      3 VIFNIDGGYLEFLTRDF 19


RESULT   6
Q9UGS1
ID   Q9UGS1      PRELIMINARY;      PRT;    12 AA.
AC   Q9UGS1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ796I17.4 (NOVEL PROTEIN SIMILAR TO GS2) (FRAGMENT).
GN   DJ796I17.4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035398; CAB63074.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   12 AA;  1430 MW;  AF7740ABECB69AA6 CRC64;


  Query Match             20.1%;  Score 27;  DB 4;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGF 13
          ::|:||  |
Db      1 MILEMDMSF 9


RESULT   7
Q9TRQ7
ID   Q9TRQ7      PRELIMINARY;      PRT;    14 AA.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:57 ; Search time 28.23 Seconds
(without alignments)
85.998 Million cell updates/sec

Title: US-08-934-367-29
Perfect score: 134
Sequence: 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 6467

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	24.6	24	6	Q9TR30	Q9tr30 ovis aries
2	31	23.1	19	6	Q9TRR6	Q9trr6 oryctolagus
3	30.5	22.8	17	4	Q9UJH1	Q9ujh1 homo sapien
4	30	22.4	21	2	P70861	P70861 borrelia bu
5	29	21.6	20	5	Q9TWH5	Q9twh5 heliothis v
6	27	20.1	12	4	Q9UGS1	Q9ugs1 homo sapien
7	27	20.1	14	6	Q9TRQ7	Q9trq7 bos taurus
8	27	20.1	22	10	Q9S8E1	Q9s8e1 capsicum an
9	26.5	19.8	20	6	Q9TRC2	Q9trc2 oryctolagus
10	26.5	19.8	21	2	Q9ZG55	Q9zg55 chlamydia t
11	26	19.4	22	6	O02830	O02830 oryctolagus
12	26	19.4	24	2	Q46081	Q46081 clostridium
13	26	19.4	25	13	P82401	P82401 litoria ran
14	25	18.7	16	3	O94554	O94554 schizosacch
15	25	18.7	19	13	Q9PS70	Q9ps70 gallus gall
16	25	18.7	20	10	Q9S878	Q9s878 petunia hyb
17	25	18.7	21	12	O93044	O93044 maize strea
18	25	18.7	21	12	O93046	O93046 maize strea
19	25	18.7	21	12	O93047	O93047 maize strea
20	25	18.7	23	6	Q9TRZ8	Q9trz8 canis famil
21	25	18.7	24	4	Q13661	Q13661 homo sapien
22	25	18.7	24	12	Q69137	Q69137 human herpe
23	24.5	18.3	22	13	Q9PS42	Q9ps42 gallus gall
24	24.5	18.3	23	12	O10423	O10423 influenza a
25	24.5	18.3	25	13	Q9PS41	Q9ps41 gallus gall
26	24	17.9	13	2	O31364	O31364 borrelia ga
27	24	17.9	13	2	O31365	O31365 borrelia ga
28	24	17.9	13	2	O34622	O34622 borrelia bu
29	24	17.9	18	1	Q9UYK7	Q9uyk7 pyrococcus
30	24	17.9	19	5	Q9W5Q8	Q9w5q8 drosophila
31	24	17.9	20	6	Q9TTG3	Q9ttg3 ateles belz
32	24	17.9	21	2	Q9X3D0	Q9x3d0 prochloroco
33	24	17.9	21	12	O93050	O93050 maize strea
34	24	17.9	22	4	Q9UQ31	Q9uq31 homo sapien
35	24	17.9	22	6	Q9XTA7	Q9xta7 cercopithec
36	24	17.9	22	6	Q9XTA6	Q9xta6 canis famil
37	24	17.9	22	11	Q9WVT2	Q9wvt2 cricetulus
38	24	17.9	22	11	Q9WVT1	Q9wvt1 rattus norv
39	24	17.9	22	11	Q9R1U8	Q9r1u8 mus musculu
40	24	17.9	26	8	O79846	O79846 asplenium n
41	24	17.9	26	10	Q9S8P2	Q9s8p2 raphanus sa
42	23.5	17.5	22	2	Q50082	Q50082 mycobacteri
43	23	17.2	8	2	O85406	O85406 coxiella bu
44	23	17.2	9	5	O96417	O96417 drosophila
45	23	17.2	10	2	Q9X534	Q9x534 leclercia a

ALIGNMENTS

RESULT 1
Q9TR30
ID Q9TR30 PRELIMINARY; PRT; 24 AA.
AC Q9TR30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT VIIA-L (FRAGMENT).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE.
RX MEDLINE; 96092035.
RA Linder D., Freund R., Kadenbach B.;
RT "Species-specific expression of cytochrome c oxidase isozymes.";
RL Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 112:461-469(1995).
SQ SEQUENCE 24 AA; 2795 MW; D49D27C03B61F803 CRC64;

Query Match 24.6%; Score 33; DB 6; Length 24;
Best Local Similarity 58.3%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       6 LLQMDFGFPEHL 17
           | | | | | ||
Db      11 LFQEDNGIPVHL 22
```

RESULT 2
Q9TRR6
ID Q9TRR6 PRELIMINARY; PRT; 19 AA.
AC Q9TRR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CALCYCLIN-ASSOCIATED PROTEIN, CAP50=CA2+/PHOSPHOLIPID-BINDING PROTEIN
DE L-14 FRAGMENT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

```
  Query Match             17.2%;  Score 23;  DB 1;  Length 25;
  Best Local Similarity   44.4%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 QMDFGFPEH 16
           : | | ||:
Db      17 ECDXGTPEN 25


RESULT  13
UC03_MAIZE
ID   UC03_MAIZE     STANDARD;      PRT;    18 AA.
AC   P80609;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 146)
DE   (FRAGMENTS).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=COLEOPTILE;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.1, ITS MW IS: 29.3 KDA.
CC   -!- CAUTION: THE ORDER OF THE PEPTIDES SHOWN IS UNCERTAIN.
DR   MAIZE-2DPAGE; P80609; COLEOPTILE.
DR   MAIZEDB; 123924; -.
FT   NON_TER       1      1
FT   NON_CONS      9     10
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1938 MW;  6F513BEBC35881C0 CRC64;


  Query Match             16.4%;  Score 22;  DB 1;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DGFL 5
           ||||
Db      13 DGFL 16


RESULT  14
CAQS_RAT
ID   CAQS_RAT       STANDARD;      PRT;    20 AA.
AC   P19633;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   CALSEQUESTRIN, SKELETAL MUSCLE ISOFORM (ASPARTACTIN) (LAMININ-BINDING
DE   PROTEIN) (FRAGMENT).
GN   CASQ1.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88331073.
RA   Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT   "Isolation and characterization of a laminin-binding protein from rat
RT   and chick muscle.";
RL   J. Cell Biol. 107:687-697(1988).
CC   -!- FUNCTION: CALSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC       CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC       IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH
CC       A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC       TO 50 MOLES OF CALCIUM. ALSO BINDS LAMININ.
CC   -!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE
CC       SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF
CC       FAST SKELETAL MUSCLE CELLS. ASPARTACTIN IS FOUND IN THE BASAL
CC       LAMINA SURROUNDING INDIVIDUAL MUSCLE FIBERS.
CC   -!- TISSUE SPECIFICITY: SKELETAL AND HEART MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
DR   PIR; A31049; A31049.
DR   INTERPRO; IPR001393; -.
DR   PROSITE; PS00864; CALSEQUESTRIN_2; PARTIAL.
DR   PROSITE; PS00863; CALSEQUESTRIN_1; 1.
KW   Muscle; Glycoprotein; Calcium-binding.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2238 MW;  92ADE04FC2A69280 CRC64;


  Query Match             16.4%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 FPEHLLVD 20
           |||:  ||
Db       6 FPEYDGVD 13


RESULT  15
MIF_PIG
ID   MIF_PIG        STANDARD;      PRT;    20 AA.
AC   P80928;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   MACROPHAGE MIGRATION INHIBITORY FACTOR (MIF) (GLYCOSYLATION-INHIBITING
DE   FACTOR) (GIF) (FRAGMENT).
GN   MIF.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE.
RA   Riviere S., Bouet F., Menez A., Galat A.;
RL   Submitted (MAR-1997) to the SWISS-PROT data bank.
CC   -!- FUNCTION: THE EXPRESSION OF MIF AT SITES OF INFLAMMATION SUGGEST
CC       A ROLE FOR THE MEDIATOR IN REGULATING THE FUNCTION OF MACROPHAGE
CC       IN HOST DEFENSE (BY SIMILARITY).
CC   -!- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MIF FAMILY.
DR   HSSP; P14174; 1MIF.
DR   INTERPRO; IPR001398; -.
DR   PROSITE; PS01158; MIF; PARTIAL.
KW   Macrophage; Inflammatory response; Cytokine.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2147 MW;  3517AF60F3012A61 CRC64;


  Query Match             16.4%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 DGFL 5
           ||||
Db      16 DGFL 19


Search completed: December 21, 2000, 08:40:18
Job time: 116 sec
```

```
RA   Linder D., Freund R., Kadenbach B.;
RT   "Species-specific expression of cytochrome c oxidase isozymes.";
RL   Comp. Biochem. Physiol. 112B:461-469(1995).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      23     23
SQ   SEQUENCE   23 AA;  2627 MW;  AD3EA34B61FF73CE CRC64;


  Query Match             18.7%;  Score 25;  DB 1;  Length 23;
  Best Local Similarity   50.0%;  Pred. No. 7.8e+02;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPEHL 17
          | | | | |  |
Db     11 LFQADNGLPVXL 22


RESULT  10
ANGT_HORSE
ID   ANGT_HORSE     STANDARD;      PRT;    14 AA.
AC   P01016;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   ANGIOTENSINOGEN (FRAGMENT).
GN   AGT.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RA   Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
RT   "The preparation, purification, and amino acid sequence of a
RT   polypeptide renin substrate.";
RL   J. Exp. Med. 106:439-453(1957).
CC   -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC       CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC       CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC       PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC       PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC       BALANCE OF BODY FLUIDS.
CC   -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   PIR; A01250; A01250.
DR   INTERPRO; IPR000215; -.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Vasoconstrictor; Plasma; Serpin.
FT   PEPTIDE       1     10       ANGIOTENSIN I.
FT   PEPTIDE       1      8       ANGIOTENSIN II.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1759 MW;  2E9921F8EEEFBDD7 CRC64;


  Query Match             17.9%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 6.6e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     14 PEHLLV 19
          | ||||
Db      7 PFHLLV 12


RESULT  11
SCX2_MESTA
ID   SCX2_MESTA     STANDARD;      PRT;    24 AA.
AC   P45668;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   NEUROTOXIN II (BT-II) (FRAGMENT).
OS   Mesobuthus tamulus (Eastern Indian scorpion) (Buthus tamulus).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Scorpiones;
OC   Buthoidea; Buthidae; Mesobuthus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 94287436.
RA   Lala K., Narayanan P.;
RT   "Purification, N-terminal sequence and structural characterization of
RT   a toxic protein from the Indian scorpion venom Buthus tamulus.";
RL   Toxicon 32:325-338(1994).
CC   -!- FUNCTION: BINDS TO SODIUM CHANNELS AND INHIBIT THE INACTIVATION OF
CC       THE ACTIVATED CHANNELS, THEREBY BLOCKING NEURONAL TRANSMISSION.
CC       THIS TOXIN IS ACTIVE AGAINST MAMMALS. LD(50) IS 2.25 MG/KG IN MICE
CC       BY SUBCUTANEOUS INJECTION.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE ALPHA/BETA-SCORPION TOXIN FAMILY.
CC       ALPHA-TOXIN SUBFAMILY.
KW   Neurotoxin; Sodium channel inhibitor.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2686 MW;  DA1990C8FF8E2769 CRC64;


  Query Match             17.2%;  Score 23;  DB 1;  Length 24;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 DGFLL 6
          ||:||
Db      2 DGYLL 6


RESULT  12
BOTR_BOTJA
ID   BOTR_BOTJA     STANDARD;      PRT;    25 AA.
AC   P22028;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   BOTROCETIN (PLATELET COAGGLUTININ) (FRAGMENT).
OS   Bothrops jararaca (Jararaca).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Crotalinae; Bothrops.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 91129280.
RA   Fujimura Y., Titani K., Usami Y., Suzuki M., Oyama R., Matsui T.,
RA   Fukui H., Sugimoto M., Ruggeri Z.M.;
RT   "Isolation and chemical characterization of two structurally and
RT   functionally distinct forms of botrocetin, the platelet coagglutinin
RT   isolated from the venom of Bothrops jararaca.";
RL   Biochemistry 30:1957-1964(1991).
CC   -!- FUNCTION: THERE ARE TWO DISTINCT FORMS OF THE VON WILLEBRAND
CC       FACTOR-DEPENDENT PLATELET COAGGLUTININ. THE DIMERIC FORM IS
CC       34-TIMES MORE ACTIVE THAN THE ONE-CHAIN BOTROCETIN IN PROMOTING
CC       VWF BINDING TO PLATELETS.
CC   -!- SUBUNIT: MONOMER. VWF AND BOTROCETIN FORM A SOLUBLE COMPLEX.
CC   -!- PTM: CONTAINS NUMEROUS INTRACHAIN DISULFIDE BONDS.
CC   -!- SIMILARITY: TO OTHER MEMBERS OF THE C-TYPE LECTIN FAMILY.
DR   INTERPRO; IPR001304; -.
DR   PROSITE; PS00615; C_TYPE_LECTIN_1; PARTIAL.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; PARTIAL.
KW   Glycoprotein; Venom.
FT   VARIANT       2      2       I -> V.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2655 MW;  D25D9031A705CAF8 CRC64;
```

```
AC   P12662;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   INTERPHOTORECEPTOR RETINOID-BINDING PROTEIN (IRBP) (INTERSTITIAL
DE   RETINOL-BINDING PROTEIN) (FRAGMENT).
GN   RBP3.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 86301171.
RA   Fong S.L., Cook R.G., Alvarez R.A., Liou G.I., Landers R.A.,
RA   Bridges C.D.B.;
RT   "N-terminal sequence homologies in interstitial retinol-binding
RT   proteins from 10 vertebrate species.";
RL   FEBS Lett. 205:309-312(1986).
CC   -!- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOIDS BETWEEN
CC       THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL
CC       PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.
CC   -!- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES
CC       THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT
CC       EPITHELIUM CELLS.
DR   PIR; B24417; B24417.
KW   Vitamin A; Transport.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2813 MW;  4E751DFA160231B7 CRC64;


  Query Match             19.4%;  Score 26;  DB 1;  Length 25;
  Best Local Similarity   38.5%;  Pred. No. 6e+02;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPEHLL 18
          :|  :: ||| |:
Db     12 ILLDNYTFPESLM 24


RESULT   7
ARCD_PSEPU
ID   ARCD_PSEPU     STANDARD;      PRT;    16 AA.
AC   P41147;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   ARGININE/ORNITHINE ANTIPORTER (FRAGMENT).
GN   ARCD.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4359;
RA   Wilson S.D., Wang M., Filpula D.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE
CC       AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC       ARGININE DEIMINASE PATHWAY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC       PERMEASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07185; AAA16963.1; -.
KW   Transport; Amino-acid transport; Transmembrane; Inner membrane.
FT   NON_TER      1      1
SQ   SEQUENCE   16 AA;  1644 MW;  90B48A7C8FAA9705 CRC64;


  Query Match             18.7%;  Score 25;  DB 1;  Length 16;
  Best Local Similarity   83.3%;  Pred. No. 5.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 DGFLLL 7
            |||| |
Db     11 DGFLTL 16


RESULT   8
COXN_THUOB
ID   COXN_THUOB     STANDARD;      PRT;    20 AA.
AC   P80980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2303 MW;  0A33BD34006E5AA6 CRC64;


  Query Match             18.7%;  Score 25;  DB 1;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 6.7e+02;
  Matches    6;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 FLLLQMDFGFPEH 16
          | |   | | | |
Db      8 FHLFYGDNGMPVH 20


RESULT   9
COXK_CANFA
ID   COXK_CANFA     STANDARD;      PRT;    23 AA.
AC   Q9TRZ8;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-HEART, MITOCHONDRIAL
DE   (EC 1.9.3.1) (CYTOCHROME C OXIDASE SUBUNIT VIIA-H) (COX VIIA-M)
DE   (FRAGMENT).
GN   COX7A1 OR COX7AH.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND HEART;
RX   MEDLINE; 96092035.
```

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 86301171.
RA   Fong S.L., Cook R.G., Alvarez R.A., Liou G.I., Landers R.A.,
RA   Bridges C.D.B.;
RT   "N-terminal sequence homologies in interstitial retinol-binding
RT   proteins from 10 vertebrate species.";
RL   FEBS Lett. 205:309-312(1986).
CC   -!- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOIDS BETWEEN
CC       THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL
CC       PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.
CC   -!- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES
CC       THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT
CC       EPITHELIUM CELLS.
DR   PIR; A24417; A24417.
KW   Vitamin A; Transport.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2799 MW;  02EDEBE61A8E4523 CRC64;


  Query Match             20.9%;  Score 28;  DB 1;  Length 24;
  Best Local Similarity   35.0%;  Pred. No. 2.8e+02;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    6;  Gaps    1;

Qy       5 LLLQM------DFGFPEHLL 18
           |:| |      :: |||:|:
Db       5 LVLDMAQVLLDNYTFPENLM 24


RESULT   3
COXJ_ONCMY
ID   COXJ_ONCMY     STANDARD;      PRT;    23 AA.
AC   P80333;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-LIVER (EC 1.9.3.1) (VIIIC)
DE   (FRAGMENT).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LIVER;
RX   MEDLINE; 94237150.
RA   Freund R., Kadenbach B.;
RT   "Identification of tissue-specific isoforms for subunits Vb and VIIa
RT   of cytochrome c oxidase isolated from rainbow trout.";
RL   Eur. J. Biochem. 221:1111-1116(1994).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
DR   PIR; S43632; S43632.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      23     23
SQ   SEQUENCE   23 AA;  2635 MW;  BCBED43FBAD9C509 CRC64;


  Query Match             20.1%;  Score 27;  DB 1;  Length 23;
  Best Local Similarity   50.0%;  Pred. No. 3.8e+02;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       6 LLQMDFGFPEHL 17
           | |   | | ||
Db      11 LFQAXNGIPVHL 22


RESULT   4
```

```
CCAA_STRTI
ID   CCAA_STRTI     STANDARD;      PRT;    24 AA.
AC   P80436;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   CHINOXALIN-2-CARBOXYLIC ACID ACTIVATING ENZYME (FRAGMENT).
OS   Streptomyces triostinicus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RA   Pahl A., Schlumbohm W., Keller U.;
RL   Submitted (MAR-1995) to the SWISS-PROT data bank.
CC   -!- FUNCTION: INVOLVED IN TRIOSTIN BIOSYNTHESIS.
KW   Antibiotic biosynthesis.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2900 MW;  91C222B657CEB6D1 CRC64;


  Query Match             20.1%;  Score 27;  DB 1;  Length 24;
  Best Local Similarity   30.0%;  Pred. No. 4e+02;
  Matches    6;  Conservative    5;  Mismatches    3;  Indels    6;  Gaps    1;

Qy       2 DGFLLLQMDFGFPEHLLVDF 21
           |||:       :|:||  ::
Db       3 DGFV------PWPDHLADEY 16


RESULT   5
YCXD_ODOSI
ID   YCXD_ODOSI     STANDARD;      PRT;    26 AA.
AC   P49839;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 3.2 KDA PROTEIN IN RPOC2-RPS2 INTERGENIC REGION (ORF26B).
OS   Odontella sinensis.
OG   Chloroplast.
OC   Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC   Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT   "The chloroplast genome of a chlorophyll a+c-containing alga,
RT   Odontella sinensis.";
RL   Plant Mol. Biol. Rep. 13:336-342(1995).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z67753; CAA91747.1; -.
KW   Chloroplast; Hypothetical protein.
SQ   SEQUENCE   26 AA;  3137 MW;  8305B587C0A30B36 CRC64;


  Query Match             20.1%;  Score 27;  DB 1;  Length 26;
  Best Local Similarity   42.9%;  Pred. No. 4.4e+02;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       2 DGFL--LLQMDFGF 13
           |||:  || ::: |
Db      11 DGFIECLLMLNYKF 24


RESULT   6
IRBP_PIG
ID   IRBP_PIG       STANDARD;      PRT;    25 AA.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:38:22 ; Search time 12.36 Seconds
(without alignments)
67.204 Million cell updates/sec

Title: US-08-934-367-29
Perfect score: 134
Sequence: 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1478

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	24.6	24	1	COXJ_SHEEP	Q9tr30 ovis aries
2	28	20.9	24	1	IRBP_SHEEP	P12663 ovis aries
3	27	20.1	23	1	COXJ_ONCMY	P80333 oncorhynchu
4	27	20.1	24	1	CCAA_STRTI	P80436 streptomyce
5	27	20.1	26	1	YCXD_ODOSI	P49839 odontella s
6	26	19.4	25	1	IRBP_PIG	P12662 sus scrofa
7	25	18.7	16	1	ARCD_PSEPU	P41147 pseudomonas
8	25	18.7	20	1	COXN_THUOB	P80980 thunnus obe
9	25	18.7	23	1	COXK_CANFA	Q9trz8 canis famil
10	24	17.9	14	1	ANGT_HORSE	P01016 equus cabal
11	23	17.2	24	1	SCX2_MESTA	P45668 mesobuthus
12	23	17.2	25	1	BOTR_BOTJA	P22028 bothrops ja
13	22	16.4	18	1	UC03_MAIZE	P80609 zea mays (m
14	22	16.4	20	1	CAQS_RAT	P19633 rattus norv
15	22	16.4	20	1	MIF_PIG	P80928 sus scrofa
16	22	16.4	24	1	FIBG_CANFA	P12800 canis famil
17	22	16.4	24	1	KPYK_CLOPA	P81344 clostridium
18	22	16.4	26	1	CATG_RAT	P17977 rattus norv
19	22	16.4	26	1	CT21_LITCI	P81847 litoria cit
20	21.5	16.0	25	1	ALR_PSEFL	P33162 pseudomonas
21	21	15.7	9	1	FARD_CALVO	P41868 calliphora
22	21	15.7	9	1	SAMP_MUSCA	P19095 mustelus ca
23	21	15.7	25	1	AMP3_MELGA	P80393 meleagris g
24	20	14.9	11	1	CH60_DROME	P35380 drosophila
25	20	14.9	11	1	TEML_RANTE	P56923 rana tempor
26	20	14.9	14	1	DCMM_PSECF	P19914 pseudomonas
27	20	14.9	16	1	MDH_SYNY4	P80460 synechocyst
28	20	14.9	21	1	ATPB_PHYPA	P80658 physcomitre
29	20	14.9	23	1	SODM_RANCA	P36215 rana catesb
30	19.5	14.6	19	1	NUO6_SOLTU	P80729 solanum tub
31	19.5	14.6	20	1	DPP4_BOVIN	P81425 bos taurus
32	19	14.2	19	1	OXLA_OPHHA	P81383 ophiophagus
33	19	14.2	19	1	UP21_UPEIN	P82027 uperoleia i
34	19	14.2	19	1	UP25_UPEIN	P82031 uperoleia i
35	19	14.2	20	1	ITRA_ALBJU	P24925 albizzia ju
36	19	14.2	20	1	SUCB_CANFA	P99507 canis famil
37	19	14.2	21	1	DCMS_PSECA	P19921 pseudomonas
38	19	14.2	21	1	YD90_HAEIN	P45194 haemophilus
39	19	14.2	23	1	CH60_THIFE	P29134 thiobacillu
40	19	14.2	25	1	ACP_ERYLO	P80919 erythrobact
41	19	14.2	25	1	CR21_LITSP	P56233 litoria spl
42	19	14.2	25	1	CR22_LITGI	P56234 litoria gil
43	19	14.2	25	1	CR23_LITCE	P56235 litoria cae
44	19	14.2	25	1	CR25_LITGI	P56237 litoria gil
45	19	14.2	25	1	FLAA_TREPH	P21983 treponema p

ALIGNMENTS

RESULT 1
COXJ_SHEEP
ID COXJ_SHEEP STANDARD; PRT; 24 AA.
AC Q9TR30;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-LIVER/HEART, MITOCHONDRIAL
DE (EC 1.9.3.1) (CYTOCHROME C OXIDASE SUBUNIT VIIA-L) (FRAGMENT).
GN COX7A2 OR COX7AL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER, AND HEART;
RX MEDLINE; 96092035.
RA Linder D., Freund R., Kadenbach B.;
RT "Species-specific expression of cytochrome c oxidase isozymes.";
RL Comp. Biochem. Physiol. 112B:461-469(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2795 MW; D49D27C03B61F803 CRC64;

Query Match 24.6%; Score 33; DB 1; Length 24;
Best Local Similarity 58.3%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 LLQMDFGFPEHL 17
| | | | | ||
Db 11 LFQEDNGIPVHL 22

RESULT 2
IRBP_SHEEP
ID IRBP_SHEEP STANDARD; PRT; 24 AA.
AC P12663;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE INTERPHOTORECEPTOR RETINOID-BINDING PROTEIN (IRBP) (INTERSTITIAL
DE RETINOL-BINDING PROTEIN) (FRAGMENT).
GN RBP3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.

```
          ||:    ||:
Db      8 DGYYYYAMDY 17


RESULT  14
S77989
cytochrome-c oxidase (EC 1.9.3.1) chain VIIb - bigeye tuna (fragment)
C;Species: Thunnus obesus (bigeye tuna)
C;Date: 17-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 30-Jan-1998
C;Accession: S77989
R;Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.
submitted to the Protein Sequence Database, June 1997
A;Reference number: S77980
A;Accession: S77989
A;Molecule type: protein
A;Residues: 1-20 <ARN>
A;Experimental source: heart
C;Genetics:
A;Genome: nuclear
C;Function:
A;Pathway: oxidative phosphorylation; respiratory chain
C;Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane


  Query Match             18.7%;  Score 25;  DB 2;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 1e+03;
  Matches    6;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      4 FLLLQMDFGFPEH 16
          | |   | | | |
Db      8 FHLFYGDNGMPVH 20


RESULT  15
PC4030
rRNA endonuclease (EC 3.1.27.10) - shiitake mushroom (fragment)
N;Alternate names: nuclease Le3
C;Species: Lentinula edodes (shiitake mushroom)
C;Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 20-Mar-1996
C;Accession: PC4030
R;Kobayashi, H.; Inokuchi, N.; Koyama, T.; Tomita, M.; Irie, M.
Biosci. Biotechnol. Biochem. 59, 1169-1171, 1995
A;Title: Purification and characterization of the 2nd 5'-nucleotide-forming nuclease fro
A;Reference number: PC4030; MUID:95337563
A;Accession: PC4030
A;Molecule type: DNA
A;Residues: 1-23 <KOB>
C;Comment: This enzyme has 3'-nucleotidase activity.
C;Keywords: endonuclease; hydrolase


  Query Match             18.7%;  Score 25;  DB 2;  Length 23;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+03;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy     11 FGFPEHLLVDFLQS 24
          :|   | || |: |
Db      1 WGMLGHELVGFIAS 14



Search completed: December 21, 2000, 08:39:30
Job time: 134 sec
```

```
A;Title: Structural and immunological studies on the soluble formate dehydrogenase from
A;Reference number: S59492; MUID:96145736
A;Accession: S59492
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRI>


  Query Match             19.4%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 5e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 QMDFGFP 14
            ::|||  |
Db       6 EIDFGTP 12


RESULT   9
B24417
interphotoreceptor retinoid-binding protein - pig (fragment)
N;Alternate names: interstitial retinol-binding protein
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C;Accession: B24417
R;Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A;Title: N-terminal sequence homologies in interstitial retinol-binding proteins from 10
A;Reference number: A91365; MUID:86301171
A;Accession: B24417
A;Molecule type: protein
A;Residues: 1-25 <FON>
C;Superfamily: interphotoreceptor retinoid-binding protein
C;Keywords: duplication


  Query Match             19.4%;  Score 26;  DB 2;  Length 25;
  Best Local Similarity   38.5%;  Pred. No. 9e+02;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       6 LLQMDFGFPEHLL 18
            :|  ::  ||| |:
Db      12 ILLDNYTFPESLM 24


RESULT  10
S71306
heat shock protein 90 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S71306
R;Conconi, M.; Szweda, L.I.; Levine, R.L.; Stadtman, E.R.; Friguet, B.
Arch. Biochem. Biophys. 331, 232-240, 1996
A;Title: Age-related decline of rat liver multicatalytic proteinase activity and protect
A;Reference number: S71306; MUID:96299287
A;Accession: S71306
A;Molecule type: protein
A;Residues: 1-15 <CON>
A;Experimental source: liver
C;Keywords: heat shock; phosphoprotein; stress-induced protein


  Query Match             18.7%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 7.2e+02;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 PEHLLVDFLQ 23
           |:| :|: |:
Db       6 PDHPIVETLR 15


RESULT  11
T44936
calmodulin kinase 2 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44936
R;Alemany, V.; Aligue, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z22873
A;Accession: T44936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-16 <ALE>
A;Cross-references: EMBL:U57982; PIDN:AAD09466.1


  Query Match             18.7%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 7.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 FGFP 14
           ||||
Db       5 FGFP 8


RESULT  12
C49048
T-cell receptor beta chain V region (CDR3 junction, clone Vbeta20.1) - human (fragmen
C;Species: Homo sapiens (man)
C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: C49048
R;Sioud, M.; Kjeldsen-Kragh, J.; Suleyman, S.; Vinje, O.; Natvig, J.B.; Forre, O.
Eur. J. Immunol. 22, 2413-2418, 1992
A;Title: Limited heterogeneity of T cell receptor variable region gene usage in juven
A;Reference number: A49048; MUID:92387250
A;Accession: C49048
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-16 <SIO>
A;Experimental source: patient EV, IL-2R+ synovial T-cells
A;Note: sequence extracted from NCBI backbone (NCBIP:113265)
C;Keywords: T-cell receptor


  Query Match             18.7%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   54.5%;  Pred. No. 7.7e+02;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 GFLLLQMDFGF 13
            || |||   |:
Db       3 GFYLLQGPPGY 13


RESULT  13
PH1607
Ig H chain V-D-J region (wild-type clone 333) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C;Accession: PH1607
R;Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A;Title: Molecular characterization of transgene-induced immunodeficiency in B-less m
A;Reference number: PH1580; MUID:93301609
A;Accession: PH1607
A;Molecule type: DNA
A;Residues: 1-17 <LEV>
A;Experimental source: bone marrow pre-B lymphocyte
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: immunoglobulin


  Query Match             18.7%;  Score 25;  DB 2;  Length 17;
  Best Local Similarity   40.0%;  Pred. No. 8.3e+02;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 DGFLLLQMDF 11
```

```
RESULT   3
B59018
MUC1 enhancer binding protein 85K chain MUC1EBP-85 - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
C;Accession: B59018
R;Abe, M.; Smith, C.J.; Larson, C.J.
submitted to the Protein Sequence Database, May 1998
A;Description: Involvement of "Ku-like" proteins in the transcription of MUC1/DF3, a bre
A;Reference number: A59018
A;Accession: B59018
A;Molecule type: protein
A;Residues: 1-17;18-26 <ABE>
A;Experimental source: breast cancer cell line MCF-7
C;Keywords: DNA binding; heterodimer


  Query Match             20.9%;  Score 28;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 4.6e+02;
  Matches    6;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       5 LLLQMDFGFPEHLLVD 20
           ::| || ||    |::
Db       9 VVLXMDVGFTLFPLIE 24


RESULT   4
A32521
hexokinase (EC 2.7.1.1) I peptide III - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C;Accession: A32521
R;Schirch, D.M.; Wilson, J.E.
Arch. Biochem. Biophys. 257, 1-12, 1987
A;Title: Rat brain hexokinase: amino acid sequence at the substrate hexose binding site
A;Reference number: A90080; MUID:87324917
A;Accession: A32521
A;Molecule type: protein
A;Residues: 1-21 <SCH>
C;Superfamily: human hexokinase I; hexokinase homology
C;Keywords: ATP; glycolysis; phosphotransferase


  Query Match             20.1%;  Score 27;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 LQMDFGFPEH 16
           |   | || |
Db       3 LGFTFSFPXH 12


RESULT   5
S43632
cytochrome-c oxidase (EC 1.9.3.1) chain VIIa, hepatic - rainbow trout (fragment)
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 20-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: S43632
R;Freund, R.; Kadenbach, B.
Eur. J. Biochem. 221, 1111-1116, 1994
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytochro
A;Reference number: S43624; MUID:94237150
A;Accession: S43632
A;Molecule type: protein
A;Residues: 1-23 <FRE>
A;Note: the source is designated as Salmo gairdneri
C;Genetics:
A;Genome: nuclear
C;Superfamily: mammalian cytochrome-c oxidase chain VIIa
C;Keywords: liver; membrane-associated complex; mitochondrion; oxidoreductase


  Query Match             20.1%;  Score 27;  DB 2;  Length 23;
  Best Local Similarity   50.0%;  Pred. No. 5.7e+02;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       6 LLQMDFGFPEHL 17
           | |   | | ||
Db      11 LFQAXNGIPVHL 22


RESULT   6
S78374
hypothetical protein 26b - Odontella sinensis chloroplast
C;Species: chloroplast Odontella sinensis
C;Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 24-Apr-1998
C;Accession: S78374
R;Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
Plant Mol. Biol. Rep. 13, 336-342, 1995
A;Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sine
A;Reference number: S78238
A;Accession: S78374
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-26 <KOW>
A;Cross-references: EMBL:Z67753; NID:g1185127; PID:e211900; PID:g1185264
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast


  Query Match             20.1%;  Score 27;  DB 2;  Length 26;
  Best Local Similarity   42.9%;  Pred. No. 6.6e+02;
  Matches    6;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       2 DGFL--LLQMDFGF 13
           |||:  || ::: |
Db      11 DGFIECLLMLNYKF 24


RESULT   7
A54077
cytochrome b558 - rabbit (fragment)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 23-Mar-1995
C;Accession: A54077
R;Escriou, V.; Laporte, F.; Garin, J.; Brandolin, G.; Vignais, P.V.
J. Biol. Chem. 269, 14007-14014, 1994
A;Title: Purification and physical properties of a novel type of cytochrome b from ra
A;Reference number: A54077; MUID:94245717
A;Accession: A54077
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <ESC>
A;Experimental source: peritoneal neutrophils
A;Note: sequence extracted from NCBI backbone (NCBIP:148739)


  Query Match             19.8%;  Score 26.5;  DB 2;  Length 20;
  Best Local Similarity   45.0%;  Pred. No. 5.8e+02;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    5;  Gaps    1;

Qy       3 GFLLLQMDFGFPEHLLVDFL 22
           |||:|  :       ||| ||
Db       6 GFLVLLV-----SALLVGFL 20


RESULT   8
S59492
formate dehydrogenase alpha chain - Alcaligenes eutrophus (fragment)
C;Species: Alcaligenes eutrophus
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59492
R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.
Biol. Chem. Hoppe-Seyler 376, 561-568, 1995
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:16 ; Search time 19.49 Seconds
(without alignments)
84.657 Million cell updates/sec

Title: US-08-934-367-29
Perfect score: 134
Sequence: 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 5255

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	20.9	19	2	A39504	octamer-binding pr
2	28	20.9	24	2	A24417	interphotoreceptor
3	28	20.9	26	2	B59018	MUC1 enhancer bind
4	27	20.1	21	2	A32521	hexokinase (EC 2.7
5	27	20.1	23	2	S43632	cytochrome-c oxida
6	27	20.1	26	2	S78374	hypothetical prote
7	26.5	19.8	20	2	A54077	cytochrome b558 -
8	26	19.4	15	2	S59492	formate dehydrogen
9	26	19.4	25	2	B24417	interphotoreceptor
10	25	18.7	15	2	S71306	heat shock protein
11	25	18.7	16	2	T44936	calmodulin kinase
12	25	18.7	16	2	C49048	T-cell receptor be
13	25	18.7	17	2	PH1607	Ig H chain V-D-J r
14	25	18.7	20	2	S77989	cytochrome-c oxida
15	25	18.7	23	2	PC4030	rRNA endonuclease
16	25	18.7	25	2	A60502	myonexin - norther
17	24.5	18.3	14	2	B61597	cytochrome P450 AL
18	24	17.9	13	2	S03879	6-phosphofructokin
19	24	17.9	14	2	A01250	angiotensin precur
20	24	17.9	15	2	A60834	angiotensin I prec
21	24	17.9	17	2	B31769	T-cell receptor de
22	24	17.9	18	2	PN0175	glutathione transf
23	24	17.9	18	2	H75063	hypothetical prote
24	24	17.9	20	2	A60822	cytochrome P450 PB
25	24	17.9	20	2	A37984	ADP,ATP carrier pr
26	23	17.2	7	2	S36662	dermorphin (Lys-7)
27	23	17.2	14	2	PA0015	seed storage prote
28	23	17.2	17	2	S71864	glutathione transf
29	23	17.2	19	2	I46654	T-cell receptor de
30	23	17.2	21	2	T07683	proteinase inhibit
31	23	17.2	23	2	I39681	exeG protein - Aer
32	23	17.2	24	2	PC2199	alicyclic amine N-
33	23	17.2	24	2	T42257	phosphoprotein pho
34	23	17.2	24	2	A53357	neurotoxin Bt-II -
35	23	17.2	25	2	S35926	T-cell receptor ga
36	23	17.2	26	2	JT0965	cytochrome-c oxida
37	23	17.2	26	2	A42218	early protein Sxl
38	22	16.4	7	1	A61324	dermorphin - Rohde
39	22	16.4	11	2	PT0250	Ig heavy chain CRD
40	22	16.4	12	2	C36201	1-aminocyclopropan
41	22	16.4	14	2	A61002	photosystem II oxy
42	22	16.4	19	2	B60822	cytochrome P450 UT
43	22	16.4	20	2	S72501	protein kinase C i
44	22	16.4	20	2	A31049	calsequestrin, fas
45	22	16.4	20	2	A47105	dystroglycan - chi

ALIGNMENTS

RESULT 1
A39504
octamer-binding protein, Ku-like, 72K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39504
R;May, G.; Sutton, C.; Gould, H.
J. Biol. Chem. 266, 3052-3059, 1991
A;Title: Purification and characterization of Ku-2, an octamer-binding protein relate
A;Reference number: A39504; MUID:91131605
A;Accession: A39504
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <MAY>

Query Match 20.9%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       7 LQMDFGF 13
            |:|| ||
Db      11 LEMDVGF 17
```

RESULT 2
A24417
interphotoreceptor retinoid-binding protein - sheep (fragment)
N;Alternate names: interstitial retinol-binding protein
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C;Accession: A24417
R;Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A;Title: N-terminal sequence homologies in interstitial retinol-binding proteins from
A;Reference number: A91365; MUID:86301171
A;Accession: A24417
A;Molecule type: protein
A;Residues: 1-24 <FON>
C;Superfamily: interphotoreceptor retinoid-binding protein
C;Keywords: duplication

Query Match 20.9%; Score 28; DB 2; Length 24;
Best Local Similarity 35.0%; Pred. No. 4.2e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 6; Gaps 1;

```
Qy       5 LLLQM------DFGFPEHLL 18
            |:| |      :: |||:|:
Db       5 LVLDMAQVLLDNYTFPENLM 24
```

```
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  401:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-401


  Query Match             20.9%;  Score 28;  DB 3;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 82;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 FPEHLL 18
           :||||:
Db       2 YPEHLV 7


RESULT  15
PCT-US91-02942-88
; Sequence 88, Application PC/TUS9102942
;  GENERAL INFORMATION:
;    APPLICANT:  ROTHLEIN, ROBERT
;    APPLICANT:  ADAIR, JOHN R
;    APPLICANT:  ATHWAL, DILJEET S
;    TITLE OF INVENTION:  HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1225 Connecticut Ave. NW Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/02942
;      FILING DATE:  19910429
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9009549.8
;      FILING DATE:  27-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FOX, SAM L
;      REGISTRATION NUMBER:  30,353
;      REFERENCE/DOCKET NUMBER:  1011.0586600
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 466-0800
;      TELEFAX:  (202) 833-8716
;  INFORMATION FOR SEQ ID NO:  88:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
PCT-US91-02942-88


  Query Match             20.9%;  Score 28;  DB 4;  Length 10;
  Best Local Similarity   55.6%;  Pred. No. 82;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GFLLLQMDF 11
           |:|||  |:
Db       2 GWLLLSFDY 10


Search completed: December 21, 2000, 08:39:07
Job time: 141 sec
```

```
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPEHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT  12
US-08-424-881-3
; Sequence 3, Application US/08424881
; Patent No. 5863782
;  GENERAL INFORMATION:
;    APPLICANT:  John Joseph HOPWOOD
;    APPLICANT:  Hamish Steele SCOTT
;    APPLICANT:  Craig Geoffrey FREEMAN
;    APPLICANT:  Charles Phillip MORRIS
;    APPLICANT:  Lianne Cheryl BLANCH
;    APPLICANT:  Xiao-Nui GUO
;    TITLE OF INVENTION:  SYNTHETIC MAMMALIAN SULPHAMIDASE
;    TITLE OF INVENTION:  AND GENETIC SEQUENCES ENCODING SAME
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,881
;      FILING DATE:  19-APR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DIGIGLIO, FRANK S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9671
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  +516 742 4343
;      TELEFAX:  +516 742 4366
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-424-881-3


  Query Match             21.6%;  Score 29;  DB 2;  Length 25;
  Best Local Similarity   53.8%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RDGFLLLQMDFGF 13
           |:  |||  | ||
Db       3 RNALLLLADDGGF 15


RESULT  13
US-08-874-763-3
; Sequence 3, Application US/08874763
; Patent No. 5972333
;  GENERAL INFORMATION:
;    APPLICANT:  John Joseph HOPWOOD
;    APPLICANT:  Hamish Steele SCOTT
;    APPLICANT:  Craig Geoffrey FREEMAN
;    APPLICANT:  Charles Phillip MORRIS
;    APPLICANT:  Lianne Cheryl BLANCH
;    APPLICANT:  Xiao-Nui GUO
;    TITLE OF INVENTION:  SYNTHETIC MAMMALIAN SULPHAMIDASE
;    TITLE OF INVENTION:  AND GENETIC SEQUENCES ENCODING SAME
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/874,763
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/424,881
;      FILING DATE:  19-APR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DIGIGLIO, FRANK S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  9671
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  +516 742 4343
;      TELEFAX:  +516 742 4366
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-874-763-3


  Query Match             21.6%;  Score 29;  DB 2;  Length 25;
  Best Local Similarity   53.8%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RDGFLLLQMDFGF 13
           |:  |||  | ||
Db       3 RNALLLLADDGGF 15


RESULT  14
US-08-159-339A-401
; Sequence 401, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
```

```
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jacobs, Seth H
;      REGISTRATION NUMBER:  32,140
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5858980e
US-08-468-540B-3


  Query Match             21.6%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 RDGFLLLQMDFG 12
           |:|||     | |
Db       5 RBGFLPRHRDTG 16


RESULT  10
US-08-827-618A-22
; Sequence 22, Application US/08827618A
; Patent No. 5998366
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,618A
;      FILING DATE:  09-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,725
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/716,909
;      FILING DATE:  18-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/586,536
;      FILING DATE:  21-SEP-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-12/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-618A-22


  Query Match             21.6%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPEHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT  11
US-08-483-952A-22
; Sequence 22, Application US/08483952A
; Patent No. 6011139
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,952A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,859
;      FILING DATE:  17-SEP-1993
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/716,909
;      FILING DATE:  18-JUN-1991
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/586,536
;      FILING DATE:  21-SEP-1990
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-7/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-952A-22


  Query Match             21.6%;  Score 29;  DB 3;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
```

```
;    FEATURE:
;      NAME/KEY:  Peptide
;      LOCATION:  1..22
;      OTHER INFORMATION:  /note= "motif C peptide from human
;      OTHER INFORMATION:  telomerase core protein 1 (TCP1)"
US-08-851-843A-170


  Query Match             22.0%;  Score 29.5;  DB 3;  Length 22;
  Best Local Similarity   40.9%;  Pred. No. 1.2e+02;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    9;  Gaps    1;

Qy       1 RDGFLLLQMDFGFPEHLLVDFL 22
           ||| ||  :|         |||
Db       2 RDGLLLRLVD---------DFL 14


RESULT   7
US-08-399-696-85
; Sequence 85, Application US/08399696
; Patent No. 5756669
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  p53-BINDING POLYPEPTIDES AND
;    TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING SAME
;    NUMBER OF SEQUENCES:  126
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,696
;      FILING DATE:  02-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/156,671
;      FILING DATE:  22-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15522-000710
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  85:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-399-696-85


  Query Match             21.6%;  Score 29;  DB 1;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      15 EHLLVDFL 22
           |||::| |
Db      12 EHLIIDGL 19


RESULT   8
US-08-484-530-22
; Sequence 22, Application US/08484530
; Patent No. 5846740
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisoc
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,530
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-8/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-530-22


  Query Match             21.6%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPEHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   9
US-08-468-540B-3
; Sequence 3, Application US/08468540B
; Patent No. 5858980
;  GENERAL INFORMATION:
;    APPLICANT:  Weiner, Howard
;    APPLICANT:  Hafler, David
;    APPLICANT:  Miller, Ariel
;    APPLICANT:  Al-Sabbagh, Ahmad
;    TITLE OF INVENTION:  SUPPRESSION OF T-CELL PROLIFERATION
;    TITLE OF INVENTION:  USING PEPTIDE FRAGMENTS OF MYELIN BASIC PROTEIN
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby P.C.
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,540B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/02942
;      FILING DATE:  19910429
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9009549.8
;      FILING DATE:  27-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FOX, SAM L
;      REGISTRATION NUMBER:  30,353
;      REFERENCE/DOCKET NUMBER:  1011.0586600
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 466-0800
;      TELEFAX:  (202) 833-8716
;  INFORMATION FOR SEQ ID NO:  98:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  AMINO ACID
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US91-02942-98


  Query Match             23.9%;  Score 32;  DB 4;  Length 26;
  Best Local Similarity   54.5%;  Pred. No. 57;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 RDGFLLLQMDF 11
           | |:|||  |:
Db      16 RGGWLLLSFDY 26


RESULT   5
US-08-199-508-2
; Sequence 2, Application US/08199508
; Patent No. 5717058
;  GENERAL INFORMATION:
;    APPLICANT:  Matthews, Maura-Ann H.
;    APPLICANT:  Stetler, Gary L.
;    APPLICANT:  Anthony-Cahill, Spencer J.
;    APPLICANT:  Anderson, David C.
;    TITLE OF INVENTION:  Modulators of Gene Expression
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  5797 Central Avenue
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  Microsoft Word 5.0a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/199,508
;      FILING DATE:  February 18, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/021,536
;      FILING DATE:  February 23, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5717058ak, Henry P.
;      REGISTRATION NUMBER:  33200
;      REFERENCE/DOCKET NUMBER:  121 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303-541-3322
;      TELEFAX:  303-444-3013
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      TOPOLOGY:  unknown to applicant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  yes
US-08-199-508-2


  Query Match             22.4%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 88;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 FGFPEHLLVDFLQ 23
           ||:| ::  | :|
Db       1 FGYPVYVFGDCVQ 13


RESULT   6
US-08-851-843A-170
; Sequence 170, Application US/08851843A
; Patent No. 6093809
;  GENERAL INFORMATION:
;    APPLICANT:  Cech, Thomas R.
;    APPLICANT:  Lingner, Joachim
;    APPLICANT:  Nakamura, Toru
;    APPLICANT:  Chapman, Karen B.
;    APPLICANT:  Morin, Gregg B.
;    APPLICANT:  Harley, Calvin
;    APPLICANT:  Andrews, William H.
;    TITLE OF INVENTION:  No. 6093809el Telomerase
;    NUMBER OF SEQUENCES:  225
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, 8th Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/851,843A
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/846,017
;      FILING DATE:  25-APR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/844,419
;      FILING DATE:  18-APR-1997
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/724,643
;      FILING DATE:  01-OCT-1996
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Apple,  Randolph T.
;      REGISTRATION NUMBER:  36,429
;      REFERENCE/DOCKET NUMBER:  015389-002930US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  170:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
```

```
  Query Match             23.9%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 47;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      15 EHLLVDFLQSLS 26
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   2
US-08-940-093-133
; Sequence 133, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,093
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0006-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6037323e
US-08-940-093-133


  Query Match             23.9%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 47;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      15 EHLLVDFLQSLS 26
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   3
US-08-940-096-133
; Sequence 133, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,096
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0005-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-133


  Query Match             23.9%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 47;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      15 EHLLVDFLQSLS 26
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   4
PCT-US91-02942-98
; Sequence 98, Application PC/TUS9102942
;  GENERAL INFORMATION:
;    APPLICANT:  ROTHLEIN, ROBERT
;    APPLICANT:  ADAIR, JOHN R
;    APPLICANT:  ATHWAL, DILJEET S
;    TITLE OF INVENTION:  HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1225 Connecticut Ave. NW Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:46 ; Search time 17.58 Seconds
(without alignments)
24.789 Million cell updates/sec

Title: US-08-934-367-29
Perfect score: 134
Sequence: 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 99815

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	23.9	22	3	US-08-940-095-133	Sequence 133, App
2	32	23.9	22	3	US-08-940-093-133	Sequence 133, App
3	32	23.9	22	3	US-08-940-096-133	Sequence 133, App
4	32	23.9	26	4	PCT-US91-02942-98	Sequence 98, Appl
5	30	22.4	20	1	US-08-199-508-2	Sequence 2, Appli
6	29.5	22.0	22	3	US-08-851-843A-170	Sequence 170, App
7	29	21.6	19	1	US-08-399-696-85	Sequence 85, Appl
8	29	21.6	20	2	US-08-484-530-22	Sequence 22, Appl
9	29	21.6	20	2	US-08-468-540B-3	Sequence 3, Appli
10	29	21.6	20	2	US-08-827-618A-22	Sequence 22, Appl
11	29	21.6	20	3	US-08-483-952A-22	Sequence 22, Appl
12	29	21.6	25	2	US-08-424-881-3	Sequence 3, Appli
13	29	21.6	25	2	US-08-874-763-3	Sequence 3, Appli
14	28	20.9	10	3	US-08-159-339A-401	Sequence 401, App
15	28	20.9	10	4	PCT-US91-02942-88	Sequence 88, Appl
16	28	20.9	22	3	US-08-940-095-134	Sequence 134, App
17	28	20.9	22	3	US-08-940-093-134	Sequence 134, App
18	28	20.9	22	3	US-08-940-096-134	Sequence 134, App
19	28	20.9	26	2	US-08-482-142-22	Sequence 22, Appl
20	28	20.9	26	2	US-08-482-142-37	Sequence 37, Appl
21	28	20.9	26	2	US-08-482-142-134	Sequence 134, App
22	28	20.9	26	2	US-08-482-142-168	Sequence 168, App
23	28	20.9	26	2	US-08-478-572-22	Sequence 22, Appl
24	28	20.9	26	2	US-08-478-572-37	Sequence 37, Appl
25	28	20.9	26	2	US-08-478-572-134	Sequence 134, App
26	28	20.9	26	2	US-08-478-572-168	Sequence 168, App
27	28	20.9	26	4	PCT-US95-04481-13	Sequence 13, Appl
28	28	20.9	26	4	PCT-US95-04481-30	Sequence 30, Appl
29	27	20.1	9	2	US-08-765-783A-83	Sequence 83, Appl
30	27	20.1	11	1	US-08-467-420A-14	Sequence 14, Appl
31	27	20.1	11	1	US-08-470-110A-14	Sequence 14, Appl
32	27	20.1	11	1	US-08-667-769A-14	Sequence 14, Appl
33	27	20.1	11	2	US-08-940-371-14	Sequence 14, Appl
34	27	20.1	11	4	PCT-US95-17082A-14	Sequence 14, Appl
35	27	20.1	15	2	US-08-482-142-85	Sequence 85, Appl
36	27	20.1	15	2	US-08-478-572-85	Sequence 85, Appl
37	27	20.1	15	3	US-08-596-257A-10	Sequence 10, Appl
38	27	20.1	15	3	US-08-860-339-10	Sequence 10, Appl
39	27	20.1	20	1	US-07-678-974D-8	Sequence 8, Appli
40	27	20.1	20	2	US-08-162-149-10	Sequence 10, Appl
41	27	20.1	20	2	US-08-945-168-13	Sequence 13, Appl
42	27	20.1	22	2	US-08-559-524A-12	Sequence 12, Appl
43	27	20.1	22	3	US-08-749-707-12	Sequence 12, Appl
44	27	20.1	26	2	US-08-482-142-84	Sequence 84, Appl
45	27	20.1	26	2	US-08-482-142-99	Sequence 99, Appl

ALIGNMENTS

```
RESULT 1
US-08-940-095-133
; Sequence 133, Application US/08940095
; Patent No. 6004925
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    APPLICANT:  Dufourcq, Jean
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,095
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0004-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6004925e
US-08-940-095-133
```

Best Local Similarity 66.7%; Pred. No. 3e+05;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       8 RGLPTG 13
           || |:|
Db       1 RGQPSG 6
```

Search completed: December 21, 2000, 08:35:33
Job time: 596 sec

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PHOSPHOINOSITIDE-SPECIFIC PHOSPHOLIPASE C ISOZYME C1 (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93280199.
RA   Min D.S., Kim D.M., Lee Y.H., Seo J., Suh P.G., Ryu S.H.;
RT   "Purification of a novel phospholipase C isozyme from bovine
RT   cerebellum.";
RL   J. Biol. Chem. 268:12207-12212(1993).
SQ   SEQUENCE   18 AA;  1927 MW;  1E37AE00CFC2AFC1 CRC64;


  Query Match             24.0%;  Score 24;  DB 6;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 GLPTGQ 14
           ||| |:
Db       7 GLPXGK 12


RESULT  12
Q9S8M8
ID   Q9S8M8      PRELIMINARY;      PRT;    19 AA.
AC   Q9S8M8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   FRIII-2-VIII-GAMMA-GLIADIN (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94284958.
RA   Fluge O., Sletten K., Fluge G., Aksnes L., Elsayed S.;
RT   "In vitro toxicity of purified gluten peptides tested by organ
RT   culture.";
RL   J. Pediatr. Gastroenterol. Nutr. 18:186-192(1994).
SQ   SEQUENCE   19 AA;  2240 MW;  FC0F56FCB41B5C81 CRC64;


  Query Match             24.0%;  Score 24;  DB 10;  Length 19;
  Best Local Similarity   66.7%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       5 PSGQVQ 10


RESULT  13
Q9TRC0
ID   Q9TRC0      PRELIMINARY;      PRT;    20 AA.
AC   Q9TRC0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ENTEROTOXIN-BINDING GLYCOPROTEIN PP16K (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94259890.
RA   Shida K., Takamizawa K., Nagaoka M., Kushiro A., Osawa T., Tsuji T.;
RT   "Enterotoxin-binding glycoproteins in a proteose-peptone fraction of
RT   heated bovine milk.";
RL   J. Dairy Sci. 77:930-939(1994).
SQ   SEQUENCE   20 AA;  2321 MW;  C738FD14F55C74C3 CRC64;


  Query Match             24.0%;  Score 24;  DB 6;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EIFQEL 6
           |:|:||
Db       7 EVFREL 12


RESULT  14
Q9PXE4
ID   Q9PXE4      PRELIMINARY;      PRT;    20 AA.
AC   Q9PXE4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   VP1 PROTEIN (FRAGMENT).
OS   Foot-and-mouth disease virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC   Aphthovirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96100820.
RA   Piatti P., Hassard S., Newman J.F., Brown F.;
RT   "Antigenic variants in a plaque-isolate of foot-and-mouth disease
RT   virus: implications for vaccine production.";
RL   Vaccine 13:781-784(1995).
SQ   SEQUENCE   20 AA;  1925 MW;  E91F87C99C19D7DB CRC64;


  Query Match             24.0%;  Score 24;  DB 12;  Length 20;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 LSRGLP 11
           ::||||
Db      15 VARGLP 20


RESULT  15
O08979
ID   O08979      PRELIMINARY;      PRT;     9 AA.
AC   O08979;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA PROVIRUS INTEGRATION SITE IN TUMOR INDUCED BY MURINE LEUKEMIA
DE   VIRUS SL3-3, ISOLATE GTT TUMOR03-3 (FRAGMENT).
GN   AML1.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NMRI; TISSUE=MURINE RETROVIRUS INDUCED TUMOR;
RX   MEDLINE; 97332339.
RA   Amtoft H.W., Sorensen A.B., Bareil C., Schmidt J., Luz A.,
RA   Pedersen F.S.;
RT   "Stability of AML1 (core) site enhancer mutations in T lymphomas
RT   induced by attenuated SL3-3 murine leukemia virus mutants.";
RL   J. Virol. 71:5080-5087(1997).
DR   EMBL; Y11802; CAA72496.1; -.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  981 MW;  293E01E865A776D8 CRC64;


  Query Match             23.0%;  Score 23;  DB 11;  Length 9;
```

```
  Query Match             24.0%;  Score 24;  DB 10;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :||| | :
Db       1 MPTGGAAI 8


RESULT   7
Q9S8V3
ID   Q9S8V3      PRELIMINARY;      PRT;    15 AA.
AC   Q9S8V3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GAMMA 1 GLIADIN (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93009000.
RA   Sjostrom H., Friis S.U., Noren O., Anthonsen D.;
RT   "Purification and characterisation of antigenic gliadins in coeliac
RT   disease.";
RL   Clin. Chim. Acta 207:227-237(1992).
SQ   SEQUENCE   15 AA;  1664 MW;  2B5C932C3CCDA72E CRC64;


  Query Match             24.0%;  Score 24;  DB 10;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       6 PSGQVQ 11


RESULT   8
Q9S8V2
ID   Q9S8V2      PRELIMINARY;      PRT;    15 AA.
AC   Q9S8V2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GAMMA 2 GLIADIN (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93009000.
RA   Sjostrom H., Friis S.U., Noren O., Anthonsen D.;
RT   "Purification and characterisation of antigenic gliadins in coeliac
RT   disease.";
RL   Clin. Chim. Acta 207:227-237(1992).
SQ   SEQUENCE   15 AA;  1742 MW;  2B5C8365ACC8B32E CRC64;


  Query Match             24.0%;  Score 24;  DB 10;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       6 PSGQVQ 11


RESULT   9
Q9QUZ5
ID   Q9QUZ5      PRELIMINARY;      PRT;    15 AA.
AC   Q9QUZ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   38-42 KDA COSTIMULATORY FACTOR (FRAGMENT).
OS   Mus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96007483.
RA   Vinay D.S., Raje M., Verma R.K., Mishra G.C.;
RT   "Characterization of novel costimulatory molecules. A protein of 38-42
RT   kDa from B cell surface is concerned with T cell activation and
RT   differentiation.";
RL   J. Biol. Chem. 270:23429-23436(1995).
SQ   SEQUENCE   15 AA;  1758 MW;  F10C664C976A5D19 CRC64;


  Query Match             24.0%;  Score 24;  DB 11;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 EIFQEL 6
           |:|:||
Db      10 ELFEEL 15


RESULT  10
Q9UQG8
ID   Q9UQG8      PRELIMINARY;      PRT;    18 AA.
AC   Q9UQG8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ENDOPLASMIC RETICULUM-GOLGI INTERMEDIATE COMPARTMENT PROTEIN ERGIC-53
DE   (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 99192442.
RA   Nichols W.C., Terry V.H., Wheatley M.A., Yang A., Zivelin A.,
RA   Ciavarella N., Stefanile C., Matsushita T., Saito H., de Bosch N.B.,
RA   Ruiz-Saez A., Torres A., Thompson A.R., Feinstein D.I., White G.C.,
RA   Negrier C., Vinciguerra C., Aktan M., Kaufman R.J., Ginsburg D.,
RA   Seligsohn U.;
RT   "ERGIC-53 gene structure and mutation analysis in 19 combined factors
RT   V and VIII deficiency families.";
RL   Blood 93:2261-2266(1999).
DR   EMBL; AF081879; AAD32486.1; -.
DR   EMBL; AF081878; AAD32486.1; JOINED.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2070 MW;  34CC39BA56F8B53B CRC64;


  Query Match             24.0%;  Score 24;  DB 4;  Length 18;
  Best Local Similarity   42.9%;  Pred. No. 1.9e+03;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 ELSRGLPTGQAQVA 18
           |  :| |  | | | |
Db       1 EFQKGHPDLQGQPA 14


RESULT  11
Q9TRG0
ID   Q9TRG0      PRELIMINARY;      PRT;    18 AA.
AC   Q9TRG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
```

```
RESULT   2
Q9S8W0
ID   Q9S8W0      PRELIMINARY;      PRT;    19 AA.
AC   Q9S8W0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   HIGH-MOLECULAR-WEIGHT GLUTENIN SUBUNIT 7 (FRAGMENT).
OS   Triticum aestivum (Wheat).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Triticum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93003354.
RA   Tao H.P., Adalsteins A.E., Kasarda D.D.;
RT   "Intermolecular disulfide bonds link specific high-molecular-weight
RT   glutenin subunits in wheat endosperm.";
RL   Biochim. Biophys. Acta 1159:13-21(1992).
SQ   SEQUENCE   19 AA;  2011 MW;  95880C10396419DA CRC64;


  Query Match             27.0%;  Score 27;  DB 10;  Length 19;
  Best Local Similarity   38.5%;  Pred. No. 6.4e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 QELSRGLPTGQAQ 16
           |:  :|  :|| |
Db       1 QQAGQGQQSGQGQ 13


RESULT   3
Q9TR96
ID   Q9TR96      PRELIMINARY;      PRT;    17 AA.
AC   Q9TR96;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   MEGAPOIETIN=MEGAKARYOCYTE GROWTH AND PLATELET PRODUCTION REGULATOR.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95062214.
RA   Kuter D.J., Beeler D.L., Rosenberg R.D.;
RT   "The purification of megapoietin: a physiological regulator of
RT   megakaryocyte growth and platelet production.";
RL   Proc. Natl. Acad. Sci. U.S.A. 91:11104-11108(1994).
SQ   SEQUENCE   17 AA;  1975 MW;  DA3A5E835A755C9D CRC64;


  Query Match             26.0%;  Score 26;  DB 6;  Length 17;
  Best Local Similarity   71.4%;  Pred. No. 8.3e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 FQELSRG 9
           ||:| ||
Db      10 FQQLLRG 16


RESULT   4
Q9R514
ID   Q9R514      PRELIMINARY;      PRT;    16 AA.
AC   Q9R514;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   17 KDA LYSINE-SPECIFIC CYSTEINE PROTEINASE (FRAGMENT).
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; CFB group; Bacteroidaceae; Porphyromonas.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94103245.
RA   Pike R., McGraw W., Potempa J., Travis J.;
RT   "Lysine- and arginine-specific proteinases from Porphyromonas
RT   gingivalis. Isolation, characterization, and evidence for the
RT   existence of complexes with hemagglutinins.";
RL   J. Biol. Chem. 269:406-411(1994).
SQ   SEQUENCE   16 AA;  1819 MW;  D864F9BF367828C6 CRC64;


  Query Match             25.0%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   46.2%;  Pred. No. 1.1e+03;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       1 EIFQELSRGLPTG 13
           |||:::   || |
Db       5 EIFRQVD--LPAG 15


RESULT   5
Q28830
ID   Q28830      PRELIMINARY;      PRT;    20 AA.
AC   Q28830;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   FRUCTOSE 6-PHOSPHATE, 2-KINASE:FRUCTOSE 2,6-BISPHOSPHATASE
DE   (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92062062.
RA   Sakata J., Uyeda K.;
RT   "Characterization of two isozymic forms of heart fructose 6-phosphate,
RT   2-kinase:fructose 2,6-bisphosphatase.";
RL   Biochem. Biophys. Res. Commun. 180:470-474(1991).
DR   EMBL; S62278; AAB20157.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2231 MW;  BD8F37CBB2470660 CRC64;


  Query Match             25.0%;  Score 25;  DB 6;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQAQVAVH 20
           |  ||  :   | |
Db       6 RDKPTAETSRAAH 18


RESULT   6
Q9S8Z2
ID   Q9S8Z2      PRELIMINARY;      PRT;    15 AA.
AC   Q9S8Z2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PHOTOSYSTEM I REACTION CENTER SUBUNIT II, PSI-D.
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Caryophyllidae; Caryophyllales;
OC   Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92249324.
RA   Lagoutte B., Vallon O.;
RT   "Purification and membrane topology of PSI-D and PSI-E, two subunits
RT   of the photosystem I reaction center.";
RL   Eur. J. Biochem. 205:1175-1185(1992).
SQ   SEQUENCE   15 AA;  1515 MW;  1A7105A4A04549A9 CRC64;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:25:37 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-2
Perfect score: 100
Sequence: 1 EIFQELSRGLPTGQAQVAVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	29.0	15	8	Q9T355	Q9t355 synagrops b
2	27	27.0	19	10	Q9S8W0	Q9s8w0 triticum ae
3	26	26.0	17	6	Q9TR96	Q9tr96 ovis aries
4	25	25.0	16	2	Q9R514	Q9r514 porphyromon
5	25	25.0	20	6	Q28830	Q28830 bos taurus
6	24	24.0	15	10	Q9S8Z2	Q9s8z2 spinacia ol
7	24	24.0	15	10	Q9S8V3	Q9s8v3 triticum ae
8	24	24.0	15	10	Q9S8V2	Q9s8v2 triticum ae
9	24	24.0	15	11	Q9QUZ5	Q9quz5 mus sp. 38-
10	24	24.0	18	4	Q9UQG8	Q9uqg8 homo sapien
11	24	24.0	18	6	Q9TRG0	Q9trg0 bos taurus
12	24	24.0	19	10	Q9S8M8	Q9s8m8 triticum ae
13	24	24.0	20	6	Q9TRC0	Q9trc0 bos taurus
14	24	24.0	20	12	Q9PXE4	Q9pxe4 foot-and-mo
15	23	23.0	9	11	O08979	O08979 mus musculu
16	23	23.0	15	5	Q9TWF5	Q9twf5 artemia (br
17	23	23.0	19	8	Q9TMN3	Q9tmn3 begonia for
18	23	23.0	20	6	Q9TRV8	Q9trv8 oryctolagus
19	22	22.0	12	2	Q02128	Q02128 desulfovibr
20	22	22.0	12	10	P82328	P82328 pisum sativ
21	22	22.0	12	11	Q9QVK4	Q9qvk4 rattus sp.
22	22	22.0	15	4	O00604	O00604 homo sapien
23	22	22.0	15	12	Q88954	Q88954 vaccinia vi
24	22	22.0	17	2	Q9R512	Q9r512 porphyromon
25	22	22.0	19	4	Q07603	Q07603 homo sapien
26	22	22.0	19	8	Q9TMN1	Q9tmn1 begonia tai
27	22	22.0	19	10	Q9S956	Q9s956 zea mays (m
28	22	22.0	19	11	Q9QV31	Q9qv31 rattus sp.
29	22	22.0	19	12	Q84274	Q84274 human papil
30	22	22.0	20	2	Q9R5R7	Q9r5r7 mycobacteri
31	22	22.0	20	13	Q9PRM4	Q9prm4 gallus gall
32	21	21.0	9	12	Q88953	Q88953 vaccinia vi
33	21	21.0	10	4	Q9UGM7	Q9ugm7 homo sapien
34	21	21.0	13	7	Q9TNQ8	Q9tnq8 homo sapien
35	21	21.0	15	4	Q9UMZ6	Q9umz6 homo sapien
36	21	21.0	16	6	Q9TR09	Q9tr09 bos taurus
37	21	21.0	16	10	Q40656	Q40656 oryza sativ
38	21	21.0	19	2	Q52033	Q52033 pseudomonas
39	21	21.0	19	4	Q9UCG2	Q9ucg2 homo sapien
40	21	21.0	19	8	Q9TMN5	Q9tmn5 begonia for
41	21	21.0	19	8	Q9TMN2	Q9tmn2 begonia for
42	21	21.0	19	8	Q9T304	Q9t304 begonia tai
43	21	21.0	19	8	Q9T303	Q9t303 begonia apt
44	21	21.0	19	10	Q9S8E2	Q9s8e2 spinacia ol
45	21	21.0	19	11	Q62637	Q62637 rattus norv

ALIGNMENTS

RESULT 1
Q9T355
ID Q9T355 PRELIMINARY; PRT; 15 AA.
AC Q9T355;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 6 (FRAGMENT).
OS Synagrops bellus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC Percomorpha; Perciformes; Percoidei; Acropomatidae; Synagrops.
RN [1]
RP SEQUENCE FROM N.A.
RA Donaldson K.A., Wilson R.R.Jr.;
RT "Amphi-Panamic geminates of snook (Percoidei:Centropomidae) provide a
RT calibration of the divergence rate in the mitochondrial DNA control
RT region of fishes.";
RL Mol. Phylogenet. Evol. 0:0-0(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Donaldson K.A., Wilson R.R. Jr.;
RT "Amphi-Panamic Geminates of Snook (Percoidei: Centropomidae) Provide a
RT Calibration of the Divergence Rate in the Mitochondrial DNA Control
RT Region of Fishes.";
RL Mol. Phylogenet. Evol. 13:208-213(1999).
DR EMBL; AF123592; AAF19093.1; -.
DR EMBL; AF155907; AAF08478.1; -.
KW Mitochondrion.
FT NON_TER 1 1
SQ SEQUENCE 15 AA; 1626 MW; D9E618DD6B812B71 CRC64;

Query Match 29.0%; Score 29; DB 8; Length 15;
Best Local Similarity 66.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ELSRGLPTG 13
||:||| |
Db 2 ELTRGLARG 10

Search completed: December 21, 2000, 08:32:52
Job time: 435 sec

```
RL   Biochem. J. 288:831-837(1992).
CC   -!- FUNCTION: DEGRADATION OF LYSOPHOSPHOLIPIDS. MAY PLAY AN IMPORTANT
CC       ROLE IN PROTECTING THESE CELLS FROM THE CYTOLYTIC EFFECTS OF THE
CC       LYSOPHOSPHOLIPIDS PRODUCED BY THE ACTIVATION OF PHOSPHOLIPASE A2.
CC   -!- CATALYTIC ACTIVITY: 2-LYSOPHOSPHATIDIYLCHOLINE + H(2)O =
CC       GLYCEROPHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- INDUCTION: INCREASED BY DIFFERENTIATION OF THE CELLS.
CC   -!- PTM: THE N-TERMINUS IS BLOCKED.
KW   Hydrolase; Lipid degradation.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2138 MW;  15B38F6D23D89567 CRC64;


  Query Match             21.0%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+03;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 SRGLPTGQA 15
           |  |||  |
Db       3 SISLPTSNA 11


RESULT  13
UC25_MAIZE
ID   UC25_MAIZE     STANDARD;      PRT;    15 AA.
AC   P80631;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 77)
DE   (FRAGMENT).
OS   Zea mays (Maize).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=COLEOPTILE;
RA   Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA   Pernollet J.-C., Zivy M., de Vienne D.;
RT   "The maize two dimensional gel protein database: towards an integrated
RT   genome analysis program.";
RL   Theor. Appl. Genet. 93:997-1005(1996).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 4.9, ITS MW IS: 31.6 KDA.
DR   MAIZE-2DPAGE; P80631; COLEOPTILE.
DR   MAIZEDB; 123957; -.
FT   NON_TER       1      1
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1580 MW;  83C54CF0CE1614D0 CRC64;


  Query Match             20.0%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 LSRGLP 11
           |||  |
Db       5 LSRSFP 10


RESULT  14
CERB_RAT
ID   CERB_RAT       STANDARD;      PRT;    16 AA.
AC   P23436; P02682;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CEREBELLIN.
OS   Rattus norvegicus (Rat), and Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=RAT;
RA   Slemmon J.R., Blacher R., Danho W., Hempstead J.L., Morgan J.I.;
RT   "Isolation and sequencing of two cerebellum-specific peptides.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:6866-6870(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=PIG; TISSUE=BRAIN;
RX   MEDLINE; 89341798.
RA   Yiangou Y., Burnet P., Nikou G., Chrysanthou B.J., Bloom S.R.;
RT   "Purification and characterisation of cerebellins from human and
RT   porcine cerebellum.";
RL   J. Neurochem. 53:886-889(1989).
CC   -!- FUNCTION: CEREBELLIN EXERTS NEUROMODULATORY FUNCTIONS. DIRECTLY
CC       STIMULATES NOREPINEPHRINE RELEASE VIA THE ADENYLATE CYCLASE/PKA-
CC       DEPENDENT SIGNALING PATHWAY; AND INDIRECTLY ENHANCES
CC       ADRENOCORTICAL SECRETION IN VIVO, THROUGH A PARACRINE MECHANISM
CC       INVOLVING MEDULLARY CATECHOLAMINE RELEASE.
CC   -!- TISSUE SPECIFICITY: LOCALIZED IN THE PURKINJE CELLS.
DR   PIR; A03135; CQRT.
DR   PIR; PL0124; PL0124.
KW   Synaptosome.
SQ   SEQUENCE   16 AA;  1633 MW;  3EFA16635343D518 CRC64;


  Query Match             20.0%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 TGQAQVA 18
           :| |:||
Db       1 SGSAKVA 7


RESULT  15
LEC3_MACPO
ID   LEC3_MACPO     STANDARD;      PRT;    20 AA.
AC   P18677;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-AUG-1991 (Rel. 19, Last annotation update)
DE   AGGLUTININ BETA-3 CHAIN (MPA).
OS   Maclura pomifera (Osage orange).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Rosales; Moraceae; Maclura.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SEED;
RX   MEDLINE; 89206218.
RA   Young N.M., Johnston R.A.Z., Szabo A.G., Watson D.C.;
RT   "Homology of the D-galactose-specific lectins from Artocarpus
RT   integrifolia and Maclura pomifera and the role of an unusual small
RT   polypeptide subunit.";
RL   Arch. Biochem. Biophys. 270:596-603(1989).
CC   -!- FUNCTION: D-GALACTOSE-SPECIFIC LECTIN, BINDS THE T-ANTIGEN
CC       STRUCTURE GAL-BETA1-3-GALNAC.
CC   -!- SUBUNIT: FORMED OF FOUR ALPHA CHAINS AND FOUR BETA CHAINS.
CC   -!- SIMILARITY: TO THE OTHER AGGLUTININ BETA CHAINS.
DR   PIR; S03988; S03988.
KW   Lectin.
SQ   SEQUENCE   20 AA;  2082 MW;  AA38811BBD6370E0 CRC64;


  Query Match             20.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 3.1e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           | |::|
Db       2 PNGKSQ 7
```

```
DE   (SUCCINYL-COA SYNTHETASE, BETA CHAIN) (SCS-BETA) (FRAGMENT).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- CATALYTIC ACTIVITY: SUCCINATE + COA + GTP = SUCCINYL-COA + GDP +
CC       ORTHOPHOSPHATE.
CC   -!- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC       ACID CYCLE.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC       OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
DR   HSC-2DPAGE; P99507; DOG.
DR   INTERPRO; IPR000303; -.
DR   PROSITE; PS01217; SUCCINYL_COA_LIG_3; PARTIAL.
KW   Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2248 MW;  BE8AEFD54DBDAC2E CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   36.4%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 EIFQELSRGLP 11
           |: ||    :|
Db      10 ELLQEAGVSIP 20


RESULT  10
RS13_PARDE
ID   RS13_PARDE     STANDARD;      PRT;     17 AA.
AC   P72180;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   30S RIBOSOMAL PROTEIN S13 (FRAGMENT).
GN   RPSM.
OS   Paracoccus denitrificans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Paracoccus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=71.11T;
RA   Perrier V., Boussac A., Meier O., Barzu O., Gilles A.M.;
RT   "Adenylate kinase from P. denitrificans, an iron and zinc binding
RT   protein, catalyzes phosphorylation of AMP and reduction of
RT   cytochrome C.";
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INVOLVED IN THE BINDING OF FMET-TRNA AND, HENCE, IN THE
CC       INITIATION OF TRANSLATION (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE S13P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U64204; AAB06329.1; ALT_INIT.
DR   INTERPRO; IPR001892; -.
DR   PROSITE; PS00646; RIBOSOMAL_S13; PARTIAL.
KW   Ribosomal protein.
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1793 MW;  43F7F60DC8670B7D CRC64;


  Query Match             21.0%;  Score 21;  DB 1;  Length 17;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 LPTGQ 14
           :|||:
Db       9 IPTGK 13


RESULT  11
CUDP_VERCH
ID   CUDP_VERCH     STANDARD;      PRT;     20 AA.
AC   P80406;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CUTICLE-DEGRADING PROTEASE-LIKE PROTEIN (EC 3.4.21.-) (CHYMOELASTASE)
DE   (FRAGMENT).
OS   Verticillium chlamydosporium.
OC   Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Diheterospora.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VC10;
RX   MEDLINE; 95247009.
RA   Segers R., Butt T.M., Keen J.N., Kerry B.R., Peberdy J.F.;
RT   "The subtilisins of the invertebrate mycopathogens Verticillium
RT   chlamydosporium and Metarhizium anisopliae are serologically and
RT   functionally related.";
RL   FEMS Microbiol. Lett. 126:227-231(1995).
CC   -!- FUNCTION: CAPABLE OF BREACHING THE INSECT CUTICLE.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC       SUBTILASE FAMILY.
DR   INTERPRO; IPR000209; -.
DR   PROSITE; PS00136; SUBTILASE_ASP; PARTIAL.
DR   PROSITE; PS00137; SUBTILASE_HIS; PARTIAL.
DR   PROSITE; PS00138; SUBTILASE_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2113 MW;  26744EC2F7729B19 CRC64;


  Query Match             21.0%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   30.0%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQAQV 17
           :| | |  ::
Db       6 QGAPXGLGRI 15


RESULT  12
LPP2_HUMAN
ID   LPP2_HUMAN     STANDARD;      PRT;     20 AA.
AC   P56642;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   LYSOPHOSPHOLIPASE HL-60 PEAK 2 (EC 3.1.1.5) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LEUKEMIA;
RX   MEDLINE; 93111958.
RA   Garsetti D., Holtsberg F., Steiner M.R., Egan R.W., Clark M.A.;
RT   "Butyric acid-induced differentiation of HL-60 cells increases the
RT   expression of a single lysophospholipase.";
```

```
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   ANNETOCIN.
OS   Eisenia foetida (Common brandling worm) (Common dung-worm).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Oligochaeta; Haplotaxida;
OC   Lumbricina; Lumbricidae; Eisenia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PITUITARY;
RX   MEDLINE; 94121660.
RA   Oumi T., Ukena K., Matsushima O., Ikeda T., Fujita T., Minakata H.,
RA   Nomoto K.;
RT   "Annetocin: an oxytocin-related peptide isolated from the earthworm,
RT   Eisenia foetida.";
RL   Biochem. Biophys. Res. Commun. 198:393-399(1994).
CC   -!- FUNCTION: POTENTIATES SPONTANEOUS CONTRACTIONS OF THE GUT AND ALSO
CC       PULSATORY CONTRACTIONS AND BLADDER-SHAKING MOVEMENT OF THE
CC       NEPHRIDIA. MAY BE INVOLVED IN OSMOREGULATION OF THE ANIMAL THROUGH
CC       NEPHRIDIAL FUNCTION.
CC   -!- SIMILARITY: BELONGS TO THE VASOPRESSIN/OXYTOCIN FAMILY.
DR   PIR; PC2021; PC2021.
DR   INTERPRO; IPR000981; -.
DR   PFAM; PF00220; hormone4; 1.
DR   PROSITE; PS00264; NEUROHYPOPHYS_HORM; FALSE_NEG.
KW   Hormone; Amidation.
FT   DISULFID      1      6
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  996 MW;  D4EEB76EB45412C9 CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 RGLPTG 13
           |  |||
Db       4 RNCPTG 9


RESULT   6
TA13_TREME
ID   TA13_TREME     STANDARD;      PRT;    13 AA.
AC   P01370;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-NOV-1991 (Rel. 20, Last annotation update)
DE   TREMEROGEN A-13.
OS   Tremella mesenterica (Jelly fungus).
OC   Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Tremellales;
OC   Tremellaceae; Tremella.
RN   [1]
RP   SEQUENCE.
RA   Sakagami Y., Yoshida M., Isogai A., Suzuki A.;
RT   "Peptide sex hormones inducing conjugation tube formation in
RT   compatible mating-type cells of Tremella mesenterica.";
RL   Science 212:1525-1527(1981).
CC   -!- FUNCTION: TREMEROGEN A-13 IS PRODUCED BY THE A MATING-TYPE CELLS
CC       AND INDUCES FORMATION OF CONJUGATION TUBES IN A MATING-TYPE CELLS.
DR   PIR; A01641; JTJG3.
KW   Lipoprotein; Prenylation; Pheromone.
FT   LIPID        13     13       FARNESYL.
SQ   SEQUENCE   13 AA;  1204 MW;  680304A9697BA864 CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 9.6e+02;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 SRGLPTG 13
           :|| |:|
Db       5 NRGDPSG 11


RESULT   7
UN12_CLOPA
ID   UN12_CLOPA     STANDARD;      PRT;    13 AA.
AC   P81353;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   UNKNOWN PROTEIN CP 12 FROM 2D-PAGE (FRAGMENT).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=W5;
RX   MEDLINE; 98291870.
RA   Flengsrud R., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.8, ITS MW IS: 42.7 KDA.
FT   VARIANT       4      4       L -> V.
FT   VARIANT       5      5       Q -> I.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1465 MW;  70AA9E98455D405B CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   33.3%;  Pred. No. 9.6e+02;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 IFQELSRGL 10
           : ||:: |:
Db       3 VLQEINPGI 11


RESULT   8
BULB_NARPS
ID   BULB_NARPS     STANDARD;      PRT;    20 AA.
AC   P80554;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   BULB PROTEIN (FRAGMENT).
OS   Narcissus pseudonarcissus (Daffodil).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Asparagales; Amaryllidaceae; Narcissus.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. GOLDEN HARVEST; TISSUE=BULB;
RA   Partis M.D., Barker P., Thomas B.;
RL   Submitted (FEB-1996) to the SWISS-PROT data bank.
FT   UNSURE        2      2       OR N.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2077 MW;  76212FF3A468AD38 CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 GLPTG 13
           ||| |
Db      12 GLPPG 16


RESULT   9
SUCB_CANFA
ID   SUCB_CANFA     STANDARD;      PRT;    20 AA.
AC   P99507;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PUTATIVE SUCCINYL-COA LIGASE [GDP-FORMING] BETA-CHAIN (EC 6.2.1.4)
```

```
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86235450.
RA   Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA   Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT   "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
RT   at risk for AIDS.";
RL   Science 232:1548-1553(1986).
CC   -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC       TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC       ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC       PROMOTER.
CC   -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC   -!- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC       BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC       WAS PERINATALLY INFECTED BY HER MOTHER.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M12507; AAB12991.1; -.
DR   HIV; M12507; TAT$WMJ2.
KW   Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW   AIDS.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1467 MW;  37CC737BFEF67AA8 CRC64;


  Query Match             25.0%;  Score 25;  DB 1;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 RGLPTG 13
          || |||
Db      6 RGDPTG 11


RESULT   3
TAT_HV1Z8
ID   TAT_HV1Z8      STANDARD;      PRT;    14 AA.
AC   P12511;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN) (FRAGMENT).
GN   TAT.
OS   Human immunodeficiency virus type 1 (Z-84 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88281278.
RA   Yourno J., Josephs S.F., Reitz M.S. Jr., Zagury D., Wong-Staal F.,
RA   Gallo R.C.;
RT   "Nucleotide sequence analysis of the env gene of a new Zairian
RT   isolate of HIV-1.";
RL   AIDS Res. Hum. Retroviruses 4:165-173(1988).
CC   -!- FUNCTION: TRANSCRIPTIONAL REGULATOR THAT ACTS BY BINDING TO THE
CC       TRANS-ACTIVATING RESPONSIVE SEQUENCE (TAR) RNA ELEMENT AND
CC       ACTIVATES TRANSCRIPTION INITIATION AND/OR ELONGATION FROM THE LTR
CC       PROMOTER.
CC   -!- SUBUNIT: BINDS CYCLIN T1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLAR.
CC   -!- MISCELLANEOUS: THE Z-84 ISOLATE WAS TAKEN FROM A 54 YEAR-OLD
CC       ZAIREAN MALE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; J03653; AAA44685.1; -.
DR   HIV; J03653; TAT$JY1.
KW   Transcription regulation; Activator; RNA-binding; Nuclear protein;
KW   AIDS.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1453 MW;  37CC737BF82D7AA8 CRC64;


  Query Match             25.0%;  Score 25;  DB 1;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 3.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 RGLPTG 13
          || |||
Db      6 RGDPTG 11


RESULT   4
COXB_THUOB
ID   COXB_THUOB     STANDARD;      PRT;    20 AA.
AC   P80974;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VB (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART, AND LIVER;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.
DR   INTERPRO; IPR002124; -.
DR   PROSITE; PS00848; COX5B; PARTIAL.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2158 MW;  76F1473E1F392BD7 CRC64;


  Query Match             25.0%;  Score 25;  DB 1;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      8 RGLPTGQAQ 16
          :|:||   |
Db      4 KGIPTDDEQ 12


RESULT   5
OXYT_EISFO
ID   OXYT_EISFO     STANDARD;      PRT;     9 AA.
AC   P42998;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:25:37 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-2
Perfect score: 100
Sequence: 1 EIFQELSRGLPTGQAQVAVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	27.0	20	1	COXF_ONCMY	P80329 oncorhynchu
2	25	25.0	14	1	TAT_HV1W2	P12509 human immun
3	25	25.0	14	1	TAT_HV1Z8	P12511 human immun
4	25	25.0	20	1	COXB_THUOB	P80974 thunnus obe
5	22	22.0	9	1	OXYT_EISFO	P42998 eisenia foe
6	22	22.0	13	1	TA13_TREME	P01370 tremella me
7	22	22.0	13	1	UN12_CLOPA	P81353 clostridium
8	22	22.0	20	1	BULB_NARPS	P80554 narcissus p
9	22	22.0	20	1	SUCB_CANFA	P99507 canis famil
10	21	21.0	17	1	RS13_PARDE	P72180 paracoccus
11	21	21.0	20	1	CUDP_VERCH	P80406 verticilliu
12	21	21.0	20	1	LPP2_HUMAN	P56642 homo sapien
13	20	20.0	15	1	UC25_MAIZE	P80631 zea mays (m
14	20	20.0	16	1	CERB_RAT	P23436 rattus norv
15	20	20.0	20	1	LEC3_MACPO	P18677 maclura pom
16	19.5	19.5	20	1	YOAH_KLEAE	P56506 klebsiella
17	19	19.0	10	1	GON1_ALLMI	P37041 alligator m
18	19	19.0	16	1	DHE2_THUTH	P20016 thunnus thy
19	19	19.0	16	1	HTPG_ACICA	P81876 acinetobact
20	19	19.0	19	1	OXLA_OPHHA	P81383 ophiophagus
21	19	19.0	20	1	ALAT_PIG	P13191 sus scrofa
22	19	19.0	20	1	DNAK_MYCAV	P80462 mycobacteri
23	18	18.0	5	1	PRCT_PERAM	P01373 periplaneta
24	18	18.0	9	1	LMIP_LOCMI	P31799 locusta mig
25	18	18.0	11	1	BPPB_AGKHA	P01021 agkistrodon
26	18	18.0	13	1	ORCK_ORCLI	P37086 orconectes
27	18	18.0	14	1	SODN_STRGR	P80732 streptomyce
28	18	18.0	15	1	C1QA_RAT	P31720 rattus norv
29	18	18.0	17	1	EFG_THEAQ	Q01697 thermus aqu
30	18	18.0	17	1	RM35_YEAST	P36530 saccharomyc
31	18	18.0	19	1	COOT_SARBU	Q09148 sarcophaga
32	18	18.0	20	1	OXLA_AGKRH	P81382 agkistrodon
33	17.5	17.5	18	1	CPAX_BOVIN	P22779 bos taurus
34	17	17.0	11	1	BPP_AGKHP	P04562 agkistrodon
35	17	17.0	14	1	CRBL_VESOR	P17236 vespa orien
36	17	17.0	14	1	PSAG_CUCSA	P42049 cucumis sat
37	17	17.0	15	1	GLN2_PINPS	P81107 pinus pinas
38	17	17.0	15	1	RBS_PHYPA	P80657 physcomitre
39	17	17.0	17	1	TRP2_LEUMA	P81733 leucophaea
40	17	17.0	18	1	DRPH_UCAPU	P08871 uca pugilat
41	17	17.0	18	1	UC21_MAIZE	P80627 zea mays (m
42	17	17.0	19	1	PTHP_STRSA	P24365 streptococc
43	17	17.0	19	1	RECQ_SALTY	P40724 salmonella
44	17	17.0	19	1	SCX6_TITBA	P56610 tityus bahi
45	17	17.0	19	1	TRP3_LEUMA	P81735 leucophaea

ALIGNMENTS

```
RESULT   1
COXF_ONCMY
ID   COXF_ONCMY      STANDARD;      PRT;     20 AA.
AC   P80329;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 94237150.
RA   Freund R., Kadenbach B.;
RT   "Identification of tissue-specific isoforms for subunits Vb and VIIa
RT   of cytochrome c oxidase isolated from rainbow trout.";
RL   Eur. J. Biochem. 221:1111-1116(1994).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VB FAMILY.
DR   HSSP; P00428; 1OCC.
DR   INTERPRO; IPR002124; -.
DR   PROSITE; PS00848; COX5B; PARTIAL.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      20      20
SQ   SEQUENCE   20 AA;  2184 MW;  6A8E00CB33E92BD7 CRC64;

  Query Match             27.0%;  Score 27;  DB 1;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQAQ 16
           :|:|| : |
Db       4 KGIPTDEEQ 12


RESULT   2
TAT_HV1W2
ID   TAT_HV1W2      STANDARD;      PRT;     14 AA.
AC   P12509;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   TAT PROTEIN (TRANSACTIVATING REGULATORY PROTEIN) (FRAGMENT).
GN   TAT.
OS   Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
```

```
Qy       8 RGLPTGQ 14
            :|:|| :
Db       4 KGIPTDE 10


RESULT  14
PT0085
protein QA600027 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Feb-1996 #sequence_revision 19-Apr-1996 #text_change 06-Jun-1997
C;Accession: PT0085
R;Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.
submitted to JIPID, December 1995
A;Description: Two dimensional electrophoresis of plant proteins and standardization of
A;Reference number: PN0173
A;Accession: PT0085
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Experimental source: leaf


  Query Match             23.0%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQA 15
            :|  ||||
Db       1 KGGATGQA 8


RESULT  15
PH0758
T-cell receptor beta chain (E22) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C;Accession: PH0758
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0758
A;Molecule type: mRNA
A;Residues: 1-16 <CAS>
A;Cross-references: EMBL:X60853; NID:g50743; PIDN:CAA43243.1; PID:g50744
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor


  Query Match             23.0%;  Score 23;  DB 2;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 PTGQAQVAV 19
           |||::   |
Db       5 PTGKSNTEV 13



Search completed: December 21, 2000, 08:30:01
Job time: 265 sec
```

```
A;Molecule type: protein
A;Residues: 1-20 <ARN>
A;Experimental source: heart; liver
C;Genetics:
A;Genome: nuclear
C;Function:
A;Pathway: oxidative phosphorylation; respiratory chain
C;Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

  Query Match             25.0%;  Score 25;  DB 2;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 8.9e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQAQ 16
           :|:||   |
Db       4 KGIPTDDEQ 12


RESULT   9
A56891
gamma 1 gliadin - wheat (fragment)
C;Species: Triticum sp. (wheat)
C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 08-Dec-1995
C;Accession: A56891
R;Sjostrom, H.; Friis, S.U.; Noren, O.; Anthonsen, D.
Clin. Chim. Acta 207, 227-237, 1992
A;Title: Purification and characterisation of antigenic gliadins in coeliac disease.
A;Reference number: A56891; MUID:93009000
A;Contents: Kadett
A;Accession: A56891
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <SJO>
A;Note: sequence extracted from NCBI backbone (NCBIP:119387)

  Query Match             24.0%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 9.6e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       6 PSGQVQ 11


RESULT  10
B56891
gamma 2 gliadin - wheat (fragment)
C;Species: Triticum sp. (wheat)
C;Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 08-Dec-1995
C;Accession: B56891
R;Sjostrom, H.; Friis, S.U.; Noren, O.; Anthonsen, D.
Clin. Chim. Acta 207, 227-237, 1992
A;Title: Purification and characterisation of antigenic gliadins in coeliac disease.
A;Reference number: A56891; MUID:93009000
A;Contents: Kadett
A;Accession: B56891
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <SJO>
A;Note: sequence extracted from NCBI backbone (NCBIP:119388)

  Query Match             24.0%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   66.7%;  Pred. No. 9.6e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       6 PSGQVQ 11


RESULT  11
S52125
gamma2-gliadin P25-27 - poulard wheat
C;Species: Triticum turgidum (poulard wheat)
C;Date: 14-Jul-1995 #sequence_revision 03-Nov-1995 #text_change 16-Feb-1997
C;Accession: S52125
R;Rocher, A.; Soriano, F.; Molina, E.; Gonzalez-Limas, G.; Mendez, E.
Biochim. Biophys. Acta 1247, 143-148, 1995
A;Title: Characterization of distinct alpha- and gamma-type gliadins and low molecula
A;Reference number: S52124; MUID:95178530
A;Accession: S52125
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <ROC>
C;Superfamily: gliadin

  Query Match             24.0%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           |:|| |
Db       6 PSGQVQ 11


RESULT  12
S33867
quinaldic acid 4-oxidoreductase large chain - Serratia marcescens
C;Species: Serratia marcescens
C;Date: 02-Dec-1993 #sequence_revision 27-Feb-1997 #text_change 07-May-1999
C;Accession: S33867
R;Fetzner, S.; Lingens, F.
Biol. Chem. Hoppe-Seyler 374, 363-376, 1993
A;Title: Microbial metabolism of quinoline and related compounds. XVIII. Purification
ns 2CC-1.
A;Reference number: S33867; MUID:93363223
A;Accession: S33867
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <FET>

  Query Match             24.0%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 PTGQAQV 17
           | |||::
Db      10 PDGQAEL 16


RESULT  13
S43626
cytochrome-c oxidase (EC 1.9.3.1) chain Vb-H - trout (fragment)
C;Species: Salmo sp. (trout)
C;Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 02-Jul-1998
C;Accession: S43626
R;Freund, R.; Kadenbach, B.
Eur. J. Biochem. 221, 1111-1116, 1994
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytoc
A;Reference number: S43624; MUID:94237150
A;Accession: S43626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <FRE>
C;Keywords: electron transfer; membrane-associated complex; oxidoreductase; respirato

  Query Match             23.0%;  Score 23;  DB 2;  Length 11;
  Best Local Similarity   42.9%;  Pred. No. 1e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
RESULT   3
S43629
cytochrome-c oxidase (EC 1.9.3.1) chain Vb-L, cardiac - rainbow trout (fragment)
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 20-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 16-Jul-1999
C;Accession: S43629
R;Freund, R.; Kadenbach, B.
Eur. J. Biochem. 221, 1111-1116, 1994
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytochrc
A;Reference number: S43624; MUID:94237150
A;Accession: S43629
A;Molecule type: protein
A;Residues: 1-14 <FRE>
A;Note: the source is designated as Salmo gairdneri
C;Genetics:
A;Genome: nuclear
C;Keywords: cardiac muscle; heart; membrane-associated complex; mitochondrion; oxidoredu

  Query Match             26.0%;  Score 26;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 GLPTGQAQ 16
            |:|| : |
Db       4 GIPTXEXQ 11


RESULT   4
B60278
24K antigen - Mycobacterium bovis (fragment)
C;Species: Mycobacterium bovis
C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 18-Jun-1993
C;Accession: B60278
R;Fifis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.
Infect. Immun. 59, 800-807, 1991
A;Title: Purification and characterization of major antigens from a Mycobacterium bovis
A;Reference number: A60278; MUID:91147217
A;Accession: B60278
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <FIF>

  Query Match             26.0%;  Score 26;  DB 2;  Length 16;
  Best Local Similarity   41.7%;  Pred. No. 4.8e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 FQELSRGLPTGQ 14
            ::|  :|  |||
Db       5 YKEELKGTDTGQ 16


RESULT   5
C53113
Lys-gingipain form 2, 17K - Porphyromonas gingivalis (fragment)
N;Alternate names: lysine-specific cysteine proteinase 2, 17K
C;Species: Porphyromonas gingivalis
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 05-Apr-1995
C;Accession: C53113
R;Pike, R.; McGraw, W.; Potempa, J.; Travis, J.
J. Biol. Chem. 269, 406-411, 1994
A;Title: Lysine- and arginine-specific proteinases from Porphyromonas gingivalis. Isolat
A;Reference number: A53113; MUID:94103245
A;Accession: C53113
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <PIK>
A;Experimental source: H66
A;Note: sequence extracted from NCBI backbone (NCBIP:141693)

  Query Match             25.0%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   46.2%;  Pred. No. 7e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    2;  Gaps    1;

Qy       1 EIFQELSRGLPTG 13
            |||:::   || |
Db       5 EIFRQVD--LPAG 15


RESULT   6
A41877
LcrKc - Yersinia pseudotuberculosis (fragment)
C;Species: Yersinia pseudotuberculosis
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C;Accession: A41877
R;Rimpilainen, M.; Forsberg, A.; Wolf-Watz, H.
J. Bacteriol. 174, 3355-3363, 1992
A;Title: A novel protein, LcrQ, involved in the low-calcium response of Yersinia pseu
A;Reference number: A41877; MUID:92250432
A;Contents: YPIII, pIB102
A;Accession: A41877
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <RIM>
A;Note: sequence extracted from NCBI backbone (NCBIN:100533, NCBIP:100534)

  Query Match             25.0%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   37.5%;  Pred. No. 8e+02;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       4 QELSRGLPTGQAQVAV 19
            : ||| : |   |: |
Db       1 EALSRAISTTLGQMKV 16


RESULT   7
I46940
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC
C;Species: Bos primigenius taurus (cattle)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 31-Mar-2000
C;Accession: I46940
R;Sakata, J.; Uyeda, K.
Biochem. Biophys. Res. Commun. 180, 470-474, 1991
A;Title: Characterization of two isozymic forms of heart fructose 6-phosphate, 2-kina
A;Reference number: I46940; MUID:92062062
A;Accession: I46940
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-20 <SAK>
A;Cross-references: GB:S62278; NID:g237973; PIDN:AAB20157.1; PID:g237974
C;Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; ph
C;Keywords: phosphoric monoester hydrolase; phosphotransferase

  Query Match             25.0%;  Score 25;  DB 2;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 8.9e+02;
  Matches    5;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       8 RGLPTGQAQVAVH 20
            |  || :   | |
Db       6 RDKPTAETSRAAH 18


RESULT   8
S77983
cytochrome-c oxidase (EC 1.9.3.1) chain Vb - bigeye tuna (fragment)
C;Species: Thunnus obesus (bigeye tuna)
C;Date: 17-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 30-Jan-1998
C;Accession: S77983
R;Arnold, S.; Lee, J.; Kim, M.; Song, E.; Linder, D.; Lottspeich, F.; Kadenbach, B.
submitted to the Protein Sequence Database, June 1997
A;Reference number: S77980
A;Accession: S77983
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:25:36 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-2
Perfect score: 100
Sequence: 1 EIFQELSRGLPTGQAQVAVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	27.0	20	2	S43627	cytochrome-c oxida
2	27	27.0	20	2	PQ0537	arylhydroxamic aci
3	26	26.0	14	2	S43629	cytochrome-c oxida
4	26	26.0	16	2	B60278	24K antigen - Myco
5	25	25.0	16	2	C53113	Lys-gingipain form
6	25	25.0	18	2	A41877	LcrKc - Yersinia p
7	25	25.0	20	2	I46940	6-phosphofructo-2-
8	25	25.0	20	2	S77983	cytochrome-c oxida
9	24	24.0	15	2	A56891	gamma 1 gliadin -
10	24	24.0	15	2	B56891	gamma 2 gliadin -
11	24	24.0	18	2	S52125	gamma2-gliadin P25
12	24	24.0	20	2	S33867	quinaldic acid 4-o
13	23	23.0	11	2	S43626	cytochrome-c oxida
14	23	23.0	15	2	PT0085	protein QA600027 -
15	23	23.0	16	2	PH0758	T-cell receptor be
16	22	22.0	8	2	A35180	neutral proteinase
17	22	22.0	9	2	PC2021	oxytocin-related p
18	22	22.0	10	2	H60588	sperm-activating p
19	22	22.0	12	2	S29479	hypothetical prote
20	22	22.0	13	1	JTJG3	tremerogen a-13 -
21	22	22.0	15	2	A40634	orf19 3' of eryK -
22	22	22.0	15	2	PH1455	T-cell receptor al
23	22	22.0	15	2	S08282	cytochrome P450K-2
24	22	22.0	16	2	A53337	regulatory protein
25	22	22.0	17	2	E53113	gingipain, 44K - P
26	22	22.0	19	2	I39327	pre-T/NK cell-asso
27	22	22.0	20	2	A54083	p190/210, fatty ac
28	22	22.0	20	2	A47105	dystroglycan - chi
29	21	21.0	9	2	PH1591	Ig H chain V-D-J r
30	21	21.0	10	2	PH0926	T-cell receptor be
31	21	21.0	12	2	PH0920	T-cell receptor be
32	21	21.0	13	2	PH0799	T-cell receptor al
33	21	21.0	16	2	A41170	photosystem II 6.1
34	21	21.0	17	1	A61339	vespulakinin 1 - e
35	21	21.0	17	2	A61557	major merozoite su
36	21	21.0	20	2	S72501	protein kinase C i
37	21	21.0	20	2	S27350	lysophospholipase
38	21	21.0	20	2	PU0033	aldose 1-epimerase
39	21	21.0	20	2	S65884	translation initia
40	20	20.0	10	2	B39517	probable methionin
41	20	20.0	11	2	D56979	collagen alpha 1(I
42	20	20.0	12	2	B56049	urinary tract ston
43	20	20.0	13	2	S63492	dissimilatory sulf
44	20	20.0	14	2	PH1566	cerebrin 30 - huma
45	20	20.0	14	2	S39930	S-allele-associate

ALIGNMENTS

RESULT 1
S43627
cytochrome-c oxidase (EC 1.9.3.1) chain Vb-H - trout (fragment)
C;Species: Salmo sp. (trout)
C;Date: 19-Mar-1997 #sequence_revision 01-Aug-1997 #text_change 02-Jul-1998
C;Accession: S43627
R;Freund, R.; Kadenbach, B.
Eur. J. Biochem. 221, 1111-1116, 1994
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytoc
A;Reference number: S43624; MUID:94237150
A;Accession: S43627
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <FRE>
C;Keywords: electron transfer; membrane-associated complex; oxidoreductase; respirato

Query Match 27.0%; Score 27; DB 2; Length 20;
Best Local Similarity 44.4%; Pred. No. 4.1e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy       8 RGLPTGQAQ 16
           :|:|| : |
Db       4 KGIPTDEEQ 12
```

RESULT 2
PQ0537
arylhydroxamic acid N,O-acetyltransferase (EC 2.3.1.-) - golden hamster (fragment)
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-May-1999
C;Accession: PQ0537
R;Sone, T.; Yamaguchi, T.; Isobe, M.; Takabatake, E.; Adachi, T.; Hirano, K.; Wang, C
Chem. Pharm. Bull. 40, 2857-2859, 1992
A;Title: Purification and characterization of hamster hepatic microsomal N,O-acetyltr
A;Reference number: PQ0537; MUID:93099627
A;Accession: PQ0537
A;Molecule type: protein
A;Residues: 1-20 <SON>
A;Experimental source: liver
C;Keywords: acyltransferase

Query Match 27.0%; Score 27; DB 2; Length 20;
Best Local Similarity 46.2%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy       8 RGLPTGQAQVAVH 20
           |   ||| :  ||
Db       7 RNTHTGQVRGLVH 19
```

```
RESULT  15
US-08-458-887-7
; Sequence 7, Application US/08458887
; Patent No. 5914261
;  GENERAL INFORMATION:
;    APPLICANT:  Boulton, Teri G. et al.
;    TITLE OF INVENTION:  A FAMILY OF MAP2 PROTEIN KINASES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/458,887
;      FILING DATE:  02-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/701,544
;      FILING DATE:  16-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6526-049
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-458-887-7


  Query Match             27.0%;  Score 27;  DB 2;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 94;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 IFQELSRGLP 11
           ||||  :|   |
Db       1 IFQETARFQP 10


Search completed: December 21, 2000, 08:31:43
Job time: 366 sec
```

```
; Sequence 100, Application US/07794288D
; Patent No. 5580953
;  GENERAL INFORMATION:
;    APPLICANT:  ELISABETH ALBRECHT,
;    APPLICANT:  HOWARD JONES,
;    APPLICANT:  LAURA S.L. GAETA,
;    APPLICANT:  KATHRYN S. PRICKETT and
;    APPLICANT:  KEVIN BEAUMONT
;    TITLE OF INVENTION:  NOVEL AMYLIN ANTAGONIST
;    TITLE OF INVENTION:  PEPTIDES AND USES THEREFOR
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LYON & LYON
;      STREET:  First Interstate World
;      STREET:  Center
;      STREET:  633 West Fifth Street,
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44
;      MEDIUM TYPE:  Mb storage
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  IBM M.S. DOS (Version
;      OPERATING SYSTEM:  5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/794,288D
;      FILING DATE:  Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/794,288
;      FILING DATE:  19-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Duft, Bradford, J.
;      REGISTRATION NUMBER:  32,219
;      REFERENCE/DOCKET NUMBER:  193/153
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  100:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 AMINO ACIDS
;      TYPE:  AMINO ACID
;      TOPOLOGY:  LINEAR
;    MOLECULE TYPE:  PEPTIDE
;    FEATURE:
;      OTHER INFORMATION:
US-07-794-288D-100


  Query Match             28.0%;  Score 28;  DB 1;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 53;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 EIFQELSRGLP 11
          :: ||| :| |
Db      4 KLSQELHKGTP 14


RESULT  13
US-09-120-365-87
; Sequence 87, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/120,365
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Nicotiana
US-09-120-365-87


  Query Match             27.0%;  Score 27;  DB 3;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 80;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 EIFQELSRGLPT 12
          ||  |:  |||:
Db      3 EIEDEVVAGLPS 14


RESULT  14
US-08-463-862-7
; Sequence 7, Application US/08463862
; Patent No. 5776751
;  GENERAL INFORMATION:
;    APPLICANT:  Boulton, Teri G. et al.
;    TITLE OF INVENTION:  A FAMILY OF MAP2 PROTEIN KINASES
;    NUMBER OF SEQUENCES:  11
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,862
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/701,544
;      FILING DATE:  16-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Misrock, S. Leslie
;      REGISTRATION NUMBER:  18,872
;      REFERENCE/DOCKET NUMBER:  6526-049
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212 790-9090
;      TELEFAX:  212 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-463-862-7


  Query Match             27.0%;  Score 27;  DB 1;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 94;
  Matches    6;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 IFQELSRGLP 11
          |||| :|  |
Db      1 IFQETARFQP 10
```

```
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/106,016
;      FILING DATE:  31-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IPC-075 (IMI-040cp)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-440-861-16


  Query Match             29.0%;  Score 29;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 54;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :|||: |:
Db      12 IPTGELQI 19


RESULT  10
US-08-440-861-17
; Sequence 17, Application US/08440861
; Patent No. 5710126
;  GENERAL INFORMATION:
;    APPLICANT:  Griffith, Irwin J.
;    APPLICANT:  Kuo, Mei-Chang
;    APPLICANT:  Luqman, Mohammad
;    TITLE OF INVENTION:  T CELL EPITOPES OF RYEGRASS POLLEN
;    TITLE OF INVENTION:  ALLERGEN
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,861
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/106,016
;      FILING DATE:  31-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IPC-075 (IMI-040cp)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-440-861-17


  Query Match             29.0%;  Score 29;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 54;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :|||: |:
Db       2 IPTGELQI 9


RESULT  11
US-08-726-306A-31
; Sequence 31, Application US/08726306A
; Patent No. 5958684
;  GENERAL INFORMATION:
;    APPLICANT:  van Leeuwen, Frederik Willem
;    APPLICANT:  Burbach, Johannes Peter Henri
;    APPLICANT:  Grosveld, Franklin G.
;    TITLE OF INVENTION:  DIAGNOSIS METHOD AND REAGENTS
;    NUMBER OF SEQUENCES:  189
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Witcoff, Ltd.
;      STREET:  1 Financial Center
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  US
;      ZIP:  02111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/726,306A
;      FILING DATE:  02-Oct-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 95/20080.4
;      FILING DATE:  02-Oct-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/009,832
;      FILING DATE:  01-Jan-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Ph.D., Kathleen M.
;      REGISTRATION NUMBER:  34,380
;      REFERENCE/DOCKET NUMBER:  96,048-A (3255/00784)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 345-9100
;      TELEFAX:  (617) 345-9111
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-726-306A-31


  Query Match             29.0%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 54;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PTGQAQ 16
           ||||:|
Db      12 PTGQSQ 17


RESULT  12
US-07-794-288D-100
```

```
Db      6 IPTGELQI 13


RESULT   7
US-08-488-551B-635
; Sequence 635, Application US/08488551B
; Patent No. 6015661
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas J. Deacon
;    APPLICANT:  Dale A. McPhee
;    APPLICANT:  David Cooper
;    TITLE OF INVENTION:  NON-PATHOGENIC STRAINS OF HIV-1
;    NUMBER OF SEQUENCES:  841
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 GARDEN CITY PLAZA
;      CITY:  GARDEN CITY
;      STATE:  NEW YORK
;      COUNTRY:  U.S.A.
;      ZIP:  11530-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,551B
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PM3864 (AU)
;      FILING DATE:  14-FEB-1994
;      APPLICATION NUMBER:  PM4002 (AU)
;      FILING DATE:  21-FEB-1994
;      APPLICATION NUMBER:  PN0284 (AU)
;      FILING DATE:  23-DEC-1994
;      APPLICATION NUMBER:  US 08/388,353
;      FILING DATE:  14-FEB-1995
;      APPLICATION NUMBER:  PN3021/95
;      FILING DATE:  17-MAY-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  FRANK S. DIGIGLIO
;      REFERENCE/DOCKET NUMBER:  9606Z
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 742-4343
;      TELEFAX:  (516) 742-4366
;  INFORMATION FOR SEQ ID NO:  635:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-488-551B-635


  Query Match             29.0%;  Score 29;  DB 3;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 35;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      7 SRGLPTG 13
            ||| |||
Db      5 SRGDPTG 11


RESULT   8
US-08-634-060-42
; Sequence 42, Application US/08634060
; Patent No. 5712136
;  GENERAL INFORMATION:
;    APPLICANT:  Wickham, Thomas J.
;    APPLICANT:  Kovesdi, Imre
;    APPLICANT:  Roelvink, Petrus W.
;    TITLE OF INVENTION:  ADENOVIRAL-MEDIATED CELL TARGETING COMMANDED BY
;    TITLE OF INVENTION:  THE ADENOVIRUS PENTON BASE PROTEIN
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Leydig, Voit & Mayer, Ltd.
;      STREET:  Two Prudential Plaza, Suite 4900
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/634,060
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/303,162
;      FILING DATE:  08-SEP-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kilyk, John Jr.
;      REGISTRATION NUMBER:  30763
;      REFERENCE/DOCKET NUMBER:  71602
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5600
;      TELEFAX:  (312) 616-5700
;  INFORMATION FOR SEQ ID NO:  42:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-634-060-42


  Query Match             29.0%;  Score 29;  DB 1;  Length 19;
  Best Local Similarity   55.6%;  Pred. No. 51;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 ELSRGLPTG 13
            :| ||:| |
Db      2 KLGRGVPRG 10


RESULT   9
US-08-440-861-16
; Sequence 16, Application US/08440861
; Patent No. 5710126
;  GENERAL INFORMATION:
;    APPLICANT:  Griffith, Irwin J.
;    APPLICANT:  Kuo, Mei-Chang
;    APPLICANT:  Luqman, Mohammad
;    TITLE OF INVENTION:  T CELL EPITOPES OF RYEGRASS POLLEN
;    TITLE OF INVENTION:  ALLERGEN
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,861
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/195,947
;      FILING DATE:  14-FEB-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,060
;      FILING DATE:  14-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hohenschutz, Liza D.
;      REGISTRATION NUMBER:  33,712
;      REFERENCE/DOCKET NUMBER:  IMPH-0024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-195-947-30


  Query Match             29.0%;  Score 29;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :|||: |:
Db       6 IPTGELQI 13


RESULT   5
US-08-433-885-30
; Sequence 30, Application US/08433885
; Patent No. 5869333
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir
;    APPLICANT:  Knox, Robert B.
;    APPLICANT:  Smith, Penelope
;    APPLICANT:  Avjioglu, Asil
;    APPLICANT:  Theerakulpisut, Piyada
;    APPLICANT:  Hough, Terryn
;    APPLICANT:  Suphioglu, Cenk
;    APPLICANT:  Ong, Eng Kok
;    TITLE OF INVENTION:  Ryegrass Pollen Allergen
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lahive & Cockfield
;      STREET:  60 State Street, Suite 510
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,885
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-039C2D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-433-885-30


  Query Match             29.0%;  Score 29;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :|||: |:
Db       6 IPTGELQI 13


RESULT   6
US-08-433-908B-30
; Sequence 30, Application US/08433908B
; Patent No. 5965455
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir
;    APPLICANT:  Knox, Robert B.
;    APPLICANT:  Smith, Penelope
;    APPLICANT:  Avjioglu, Asil
;    APPLICANT:  Theerakulpisut, Piyada
;    APPLICANT:  Hough, Terryn
;    APPLICANT:  Suphioglu, Cenk
;    APPLICANT:  Ong, Eng Kok
;    TITLE OF INVENTION:  Ryegrass Pollen Allergen
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lahive & Cockfield, LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,908B
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mandragouras, Amy E.
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IMI-039C2D4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 742-4214
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-433-908B-30


  Query Match             29.0%;  Score 29;  DB 2;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
           :|||: |:
```

```
Qy      7 SRGLPTGQA 15
          ||| | ||:
Db      5 SRGSPGGQS 13


RESULT   2
US-08-433-854-30
; Sequence 30, Application US/08433854
; Patent No. 5721119
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir
;    APPLICANT:  Knox, Robert B.
;    APPLICANT:  Smith, Penelope
;    APPLICANT:  Avjioglu, Asil
;    APPLICANT:  Theerakulpisut, Piyada
;    APPLICANT:  Hough, Terryn
;    APPLICANT:  Suphioglu, Cenk
;    APPLICANT:  Ong, Eng Kok
;    TITLE OF INVENTION:  Ryegrass Pollen Allergen
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5721119ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/433,854
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,060
;      FILING DATE:  14-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hohenschutz, Liza D.
;      REGISTRATION NUMBER:  33,712
;      REFERENCE/DOCKET NUMBER:  IMPH-0024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-433-854-30


  Query Match             29.0%;  Score 29;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
             :|||: |:
Db       6 IPTGELQI 13


RESULT   3
US-08-174-745A-30
; Sequence 30, Application US/08174745A
; Patent No. 5736362
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir
;    APPLICANT:  Knox, Robert B.
;    APPLICANT:  Smith, Penelope
;    APPLICANT:  Avjioglu, Asil
;    APPLICANT:  Theerakulpisut, Piyada
;    APPLICANT:  Hough, Terryn
;    APPLICANT:  Suphioglu, Cenk
;    APPLICANT:  Ong, Eng Kok
;    TITLE OF INVENTION:  Ryegrass Pollen Allergen
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5736362ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/174,745A
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/930,060
;      FILING DATE:  14-AUG-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hohenschutz, Liza D.
;      REGISTRATION NUMBER:  33,712
;      REFERENCE/DOCKET NUMBER:  IMPH-0024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-174-745A-30


  Query Match             29.0%;  Score 29;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 32;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LPTGQAQV 17
             :|||: |:
Db       6 IPTGELQI 13


RESULT   4
US-08-195-947-30
; Sequence 30, Application US/08195947
; Patent No. 5840316
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir
;    APPLICANT:  Knox, Robert B.
;    APPLICANT:  Smith, Penelope
;    APPLICANT:  Avjioglu, Asil
;    APPLICANT:  Theerakulpisut, Piyada
;    APPLICANT:  Hough, Terryn
;    APPLICANT:  Suphioglu, Cenk
;    APPLICANT:  Ong, Eng Kok
;    TITLE OF INVENTION:  Ryegrass Pollen Allergen
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5840316ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:25:37 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-2
Perfect score: 100
Sequence: 1 EIFQELSRGLPTGQAQVAVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	30.0	16	1	US-08-418-444A-9	Sequence 9, Appli
2	29	29.0	13	1	US-08-433-854-30	Sequence 30, Appl
3	29	29.0	13	1	US-08-174-745A-30	Sequence 30, Appl
4	29	29.0	13	2	US-08-195-947-30	Sequence 30, Appl
5	29	29.0	13	2	US-08-433-885-30	Sequence 30, Appl
6	29	29.0	13	2	US-08-433-908B-30	Sequence 30, Appl
7	29	29.0	14	3	US-08-488-551B-635	Sequence 635, App
8	29	29.0	19	1	US-08-634-060-42	Sequence 42, Appl
9	29	29.0	20	1	US-08-440-861-16	Sequence 16, Appl
10	29	29.0	20	1	US-08-440-861-17	Sequence 17, Appl
11	29	29.0	20	2	US-08-726-306A-31	Sequence 31, Appl
12	28	28.0	14	1	US-07-794-288D-100	Sequence 100, App
13	27	27.0	14	3	US-09-120-365-87	Sequence 87, Appl
14	27	27.0	16	1	US-08-463-862-7	Sequence 7, Appli
15	27	27.0	16	2	US-08-458-887-7	Sequence 7, Appli
16	27	27.0	16	3	US-08-622-277A-7	Sequence 7, Appli
17	27	27.0	16	3	US-08-622-277A-12	Sequence 12, Appl
18	27	27.0	17	3	US-08-622-277A-3	Sequence 3, Appli
19	27	27.0	19	1	US-07-640-473-2	Sequence 2, Appli
20	27	27.0	19	1	US-07-908-317-21	Sequence 21, Appl
21	27	27.0	19	4	PCT-US93-06171-21	Sequence 21, Appl
22	27	27.0	20	1	US-07-678-974D-9	Sequence 9, Appli
23	27	27.0	20	2	US-08-945-168-14	Sequence 14, Appl
24	26.5	26.5	19	1	US-08-244-116B-2	Sequence 2, Appli
25	26	26.0	15	1	US-07-859-291C-10	Sequence 10, Appl
26	26	26.0	17	4	PCT-US95-04617-3	Sequence 3, Appli
27	25	25.0	9	3	US-08-159-339A-1006	Sequence 1006, Ap
28	25	25.0	10	3	US-08-159-339A-1024	Sequence 1024, Ap
29	25	25.0	11	5	5200183-7	Patent No. 5200183
30	25	25.0	11	5	5200183-10	Patent No. 5200183
31	25	25.0	11	5	5200183-11	Patent No. 5200183
32	25	25.0	12	1	US-08-433-854-21	Sequence 21, Appl
33	25	25.0	12	1	US-08-174-745A-21	Sequence 21, Appl
34	25	25.0	12	2	US-08-195-947-21	Sequence 21, Appl
35	25	25.0	12	2	US-08-433-885-21	Sequence 21, Appl
36	25	25.0	12	2	US-08-433-908B-21	Sequence 21, Appl
37	25	25.0	12	4	PCT-US93-11703-3	Sequence 3, Appli
38	25	25.0	12	4	PCT-US93-11703-13	Sequence 13, Appl
39	25	25.0	12	4	PCT-US93-11703-14	Sequence 14, Appl
40	25	25.0	12	4	PCT-US93-11703-15	Sequence 15, Appl
41	25	25.0	12	4	PCT-US93-11703-16	Sequence 16, Appl
42	25	25.0	12	4	PCT-US93-11703-17	Sequence 17, Appl
43	25	25.0	13	1	US-08-291-601-2	Sequence 2, Appli
44	25	25.0	14	1	US-08-320-373-87	Sequence 87, Appl
45	25	25.0	15	1	US-08-625-691-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT   1
US-08-418-444A-9
; Sequence 9, Application US/08418444A
; Patent No. 5773688
;  GENERAL INFORMATION:
;    APPLICANT:  KURODA, HISAO
;    APPLICANT:  HIROTA, NAOHIKO
;    APPLICANT:  ITO, KAZUTOSHI
;    TITLE OF INVENTION:  GENE EXPRESSION VECTOR USING THE GENE
;    TITLE OF INVENTION:  EXPRESSION REGULATING REGION OF THE ADP RIBOSYLATION
;    TITLE OF INVENTION:  FACTOR
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
;      STREET:  1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VIRGINIA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/418,444A
;      FILING DATE:  07-APR-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP HEI 6-71048
;      FILING DATE:  08-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Oblon, No. 5773688man F.
;      REGISTRATION NUMBER:  24,618
;      REFERENCE/DOCKET NUMBER:  2589-024-0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-418-444A-9


  Query Match           30.0%;  Score 30;  DB 1;  Length 16;
  Best Local Similarity  66.7%;  Pred. No. 27;
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRB73; TISSUE=HYPOCOTYL;
RX   MEDLINE; 95140641.
RA   Thomson M.C., Macfarlane J.L., Beagley C.T., Wolstenholme D.R.;
RT   "RNA editing of mat-r transcripts in maize and soybean increases
RT   similarity of the encoded protein to fungal and bryophyte group II
RT   intron maturases: evidence that mat-r encodes a functional protein.";
RL   Nucleic Acids Res. 22:5745-5752(1994).
DR   EMBL; U09986; AAA67712.1; -.
DR   INTERPRO; IPR001694; -.
DR   PFAM; PF00146; NADHdh; 1.
KW   Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2082 MW;  C2DA5548846188D3 CRC64;


  Query Match             21.2%;  Score 22;  DB 8;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 4e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GEKAMMLL 11
             || | |:|
Db      11 GEYANMIL 18


Search completed: December 21, 2000, 08:35:40
Job time: 603 sec
```

```
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20


RESULT  11
Q9TRQ4
ID   Q9TRQ4      PRELIMINARY;      PRT;    20 AA.
AC   Q9TRQ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN PEPTIDE L-9, CAP-50=ANNEXIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92317074.
RA   Mizutani A., Usuda N., Tokumitsu H., Minami H., Yasui K.,
RA   Kobayashi R., Hidaka H.;
RT   "CAP-50, a newly identified annexin, localizes in nuclei of cultured
RT   fibroblast 3Y1 cells.";
RL   J. Biol. Chem. 267:13498-13504(1992).
SQ   SEQUENCE   20 AA;  2296 MW;  89492C83166F4523 CRC64;


  Query Match             22.1%;  Score 23;  DB 6;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DITGE 5
           ||||:
Db       5 DITGD 9


RESULT  12
Q05403
ID   Q05403      PRELIMINARY;      PRT;     8 AA.
AC   Q05403;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FY1679;
RX   MEDLINE; 96021609.
RA   Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT   "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT   more than twice as many unknown as known open reading frames.";
RL   Yeast 11:975-986(1995).
DR   EMBL; X83121; CAA58183.1; -.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;  879 MW;  7B5322D2C441E058 CRC64;


  Query Match             21.2%;  Score 22;  DB 3;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      17 GLHN 20
           |:||
Db       1 GIHN 4


RESULT  13
```

```
Q9T2S1
ID   Q9T2S1      PRELIMINARY;      PRT;    17 AA.
AC   Q9T2S1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOCHROME-C REDUCTASE 55 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS   Solanum tuberosum (Potato).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RA   Braun H.P., Kruft V., Schmitz U.K.;
RL   Planta 193:99-106(1994).
SQ   SEQUENCE   17 AA;  1870 MW;  207804E213CD4009 CRC64;


  Query Match             21.2%;  Score 22;  DB 8;  Length 17;
  Best Local Similarity   44.4%;  Pred. No. 3.6e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 KAMMLLGQV 14
           ||: :|| :
Db       1 KAVDILGDI 9


RESULT  14
Q9S8U7
ID   Q9S8U7      PRELIMINARY;      PRT;    17 AA.
AC   Q9S8U7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SOLANIDINE UDP-GLUCOSE GLUCOSYLTRANSFERASE, SGT=19 KDA CNBR CLEAVED
DE   FRAGMENT.
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93044546.
RA   Stapleton A., Allen P.V., Tao H.P., Belknap W.R., Friedman M.;
RT   "Partial amino acid sequence of potato solanidine UDP-glucose
RT   glucosyltransferase purified by new anion-exchange and size exclusion
RT   media.";
RL   Protein Expr. Purif. 3:85-92(1992).
SQ   SEQUENCE   17 AA;  1783 MW;  75BADCB0077E593C CRC64;


  Query Match             21.2%;  Score 22;  DB 10;  Length 17;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+03;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DITGEK 6
           :|:|||
Db       8 NISGEK 13


RESULT  15
Q36277
ID   Q36277      PRELIMINARY;      PRT;    19 AA.
AC   Q36277;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT I (FRAGMENT).
GN   NAD1.
OS   Zea mays (Maize).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
```

```
DR   EMBL; AF108100; AAD16884.1; -.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   11 AA;  1203 MW;  8BADC77C6B59C33A CRC64;


  Query Match             23.1%;  Score 24;  DB 5;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 1e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLG 12
           ||:|   |:|
Db       2 TGDKFYGLMG 11


RESULT   7
Q9S739
ID   Q9S739      PRELIMINARY;      PRT;    20 AA.
AC   Q9S739;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CALMODULIN-1 (FRAGMENT).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids II; Brassicales;
OC   Brassicaceae; Arabidopsis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ECOTYPE BE-0, ECOTYPE C24, ECOTYPE LER-0, AND ECOTYPE COL-4;
RA   Loridon K., Cournoyer B., Goubely C., Depeiges A., Picard G.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF027616; AAD01853.1; -.
DR   EMBL; AF027613; AAD01850.1; -.
DR   EMBL; AF027614; AAD01851.1; -.
DR   EMBL; AF027615; AAD01852.1; -.
DR   INTERPRO; IPR002048; -.
DR   PFAM; PF00036; efhand; 1.
DR   PROSITE; PS00018; EF_HAND; UNKNOWN_1.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2303 MW;  DA1407B6ADFA2006 CRC64;


  Query Match             23.1%;  Score 24;  DB 10;  Length 20;
  Best Local Similarity   31.2%;  Pred. No. 2e+03;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    6;  Gaps    1;

Qy       1 DITGEKAMMLLGQVKY 16
           |: |:      ||: |
Db       1 DVDGD------GQINY 10


RESULT   8
Q9TQZ5
ID   Q9TQZ5      PRELIMINARY;      PRT;    17 AA.
AC   Q9TQZ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96273610.
RA   Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA   Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT   "Studies on the mechanism of early onset macular degeneration in
RT   cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT   of two proteins in the retina.";
RL   Exp. Eye Res. 62:211-219(1996).
SQ   SEQUENCE   17 AA;  1671 MW;  52CCD0D1A98B3DAF CRC64;


  Query Match             22.1%;  Score 23;  DB 6;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GQVKYGLH 19
           |:|| |::
Db       1 GKVKVGVN 8


RESULT   9
O34197
ID   O34197      PRELIMINARY;      PRT;    20 AA.
AC   O34197;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   GROES (FRAGMENT).
GN   GROES.
OS   Rickettsia rickettsii.
OC   Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R;
RX   MEDLINE; 97373904.
RA   Sumner J.W., Nicholson W.L., Massung R.F.;
RT   "PCR amplification and comparison of nucleotide sequences from the
RT   groESL heat shock operon of Ehrlichia species.";
RL   J. Clin. Microbiol. 35:2087-2092(1997).
DR   EMBL; U96733; AAB65634.1; -.
DR   INTERPRO; IPR001476; -.
DR   PFAM; PF00166; cpn10; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2235 MW;  EB6FEB9F3711EE52 CRC64;


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   25.0%;  Pred. No. 2.9e+03;
  Matches    5;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy       1 DITGEKAMMLLGQVKYGLHN 20
           :| | | :::     :|: |
Db       1 EIKGTKLIVMKESDVFGIIN 20


RESULT   10
Q9R5L1
ID   Q9R5L1      PRELIMINARY;      PRT;    20 AA.
AC   Q9R5L1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   THYMIDINE 5'-DIPHOSPHATE D-GLUCOSE 4,6-DEHYDRATASE (EC 4.2.1.46)
DE   (FRAGMENT).
OS   Streptomyces peucetius.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92268857.
RA   Thompson M.W., Strohl W.R., Floss H.G.;
RT   "Purification and characterization of TDP-D-glucose 4,6-dehydratase
RT   from anthracycline-producing streptomycetes.";
RL   J. Gen. Microbiol. 138:779-786(1992).
SQ   SEQUENCE   20 AA;  1943 MW;  7958730238C00AC9 CRC64;


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 2.9e+03;
```

```
OS   Chlorogonium elongatum.
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Haematococcaceae; Chlorogonium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2E;
RX   MEDLINE; 97306270.
RA   Kroymann J., Zetsche K.;
RT   "The apocytochrome-b gene in Chlorogonium elongatum
RT   (Chlamydomonadaceae): an intronic GIY-YIG ORF in green algal
RT   mitochondria.";
RL   Curr. Genet. 31:414-418(1997).
CC   -!- CATALYTIC ACTIVITY: NADH + ACCEPTOR = NAD(+) + REDUCED ACCEPTOR.
CC   -!- COFACTOR: FLAVOPROTEIN; IRON-SULFUR.
DR   EMBL; Y07814; CAA69147.1; -.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       1       1
SQ   SEQUENCE   17 AA;  1844 MW;  BC88B2E8E529A229 CRC64;


  Query Match             26.0%;  Score 27;  DB 8;  Length 17;
  Best Local Similarity   31.2%;  Pred. No. 5.3e+02;
  Matches    5;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLGQVKYGL 18
           ||:| : ::   |  |
Db       2 TGDKGLFIVHDQKVSL 17


RESULT   3
O31045
ID   O31045      PRELIMINARY;      PRT;    19 AA.
AC   O31045;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   LEUA LEADER PEPTIDE.
GN   LEUA.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D132;
RA   Potter C.A., Baumberg S.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF026444; AAB82585.1; -.
SQ   SEQUENCE   19 AA;  2080 MW;  1A591DC2999760D4 CRC64;


  Query Match             24.0%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           :|| :::||
Db       1 MLGIMRFGL 9


RESULT   4
Q9TWD0
ID   Q9TWD0      PRELIMINARY;      PRT;    19 AA.
AC   Q9TWD0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   SIALOPROTEIN I (FRAGMENT).
OS   Sphaerechinus granularis (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;
OC   Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;
OC   Sphaerechinus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96428846.
RA   Karamanos N.K., Manouras A., Anagnostides S., Makatsori E.,
RA   Tsegenidis T., Antonopoulos C.A.;
RT   "Isolation, biochemical and immunological characterisation of two sea
RT   urchin glycoproteins bearing sulphated poly(sialic acid)
RT   polysaccharides rich in N-glycolyl neuraminic acid.";
RL   Biochimie 78:171-182(1996).
SQ   SEQUENCE   19 AA;  2265 MW;  38FFDEB1A033E850 CRC64;


  Query Match             24.0%;  Score 25;  DB 5;  Length 19;
  Best Local Similarity   27.3%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 KAMMLLGQVKY 16
           : |:  |:::|
Db       8 QXMLTXGEIEY 18


RESULT   5
Q53545
ID   Q53545      PRELIMINARY;      PRT;    19 AA.
AC   Q53545;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   VIRF (FRAGMENT).
GN   VIRF.
OS   Shigella sonnei.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Shigella.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95394812.
RA   Nakayama S., Watanabe H.;
RT   "Involvement of cpxA, a sensor of a two-component regulatory system,
RT   in the pH-dependent regulation of expression of Shigella sonnei virF
RT   gene.";
RL   J. Bacteriol. 177:5062-5069(1995).
DR   EMBL; S79443; AAB35192.1; -.
FT   NON_TER      19      19
SQ   SEQUENCE   19 AA;  2344 MW;  52A53B9DAAEEFCF4 CRC64;


  Query Match             23.6%;  Score 24.5;  DB 2;  Length 19;
  Best Local Similarity   38.9%;  Pred. No. 1.6e+03;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    5;  Gaps    1;

Qy       8 MMLLG-----QVKYGLHN 20
           || :|      :|  |||
Db       1 MMDMGHKNKIDIKVRLHN 18


RESULT   6
Q9UAR8
ID   Q9UAR8      PRELIMINARY;      PRT;    11 AA.
AC   Q9UAR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   SIALOKININ I PREPROPROTEIN (FRAGMENT).
OS   Aedes aegypti (Yellowfever mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Culicidae; Aedes.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ROCKEFELLER/RED; TISSUE=SALIVARY GLAND;
RA   Beerntsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT   "Characterization of the sialokinin I gene encoding the salivary
RT   vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL   Insect Mol. Biol. 0:0-0(1999).
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:39 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-13
Perfect score: 104
Sequence: 1 DITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	27.9	19	10	Q9S8G6	Q9s8g6 colocasia e
2	27	26.0	17	8	O03888	O03888 chlorogoniu
3	25	24.0	19	2	O31045	O31045 streptomyce
4	25	24.0	19	5	Q9TWD0	Q9twd0 sphaerechin
5	24.5	23.6	19	2	Q53545	Q53545 shigella so
6	24	23.1	11	5	Q9UAR8	Q9uar8 aedes aegyp
7	24	23.1	20	10	Q9S739	Q9s739 arabidopsis
8	23	22.1	17	6	Q9TQZ5	Q9tqz5 macaca fasc
9	23	22.1	20	2	O34197	O34197 rickettsia
10	23	22.1	20	2	Q9R5L1	Q9r5l1 streptomyce
11	23	22.1	20	6	Q9TRQ4	Q9trq4 bos taurus
12	22	21.2	8	3	Q05403	Q05403 saccharomyc
13	22	21.2	17	8	Q9T2S1	Q9t2s1 solanum tub
14	22	21.2	17	10	Q9S8U7	Q9s8u7 solanum tub
15	22	21.2	19	8	Q36277	Q36277 zea mays (m
16	22	21.2	19	10	Q9S901	Q9s901 vigna sinen
17	22	21.2	20	4	Q9UCP6	Q9ucp6 homo sapien
18	22	21.2	20	6	Q28734	Q28734 oryctolagus
19	22	21.2	20	13	Q9PRM4	Q9prm4 gallus gall
20	21	20.2	15	6	Q9TRW0	Q9trw0 oryctolagus
21	21	20.2	17	4	Q9UCF0	Q9ucf0 homo sapien
22	21	20.2	17	11	Q9QY07	Q9qy07 mus musculu
23	21	20.2	18	4	Q9UCB2	Q9ucb2 homo sapien
24	21	20.2	19	2	Q57012	Q57012 staphylococ
25	21	20.2	19	5	Q9TWK7	Q9twk7 trypanosoma
26	21	20.2	19	11	Q64132	Q64132 rattus norv
27	21	20.2	19	11	Q9QXP9	Q9qxp9 mus musculu
28	21	20.2	20	13	Q9PSH5	Q9psh5 gallus gall
29	20	19.2	10	3	Q9UVW2	Q9uvw2 schizophyll
30	20	19.2	10	5	Q25355	Q25355 locusta mig
31	20	19.2	11	12	Q84247	Q84247 polyomaviru
32	20	19.2	14	2	Q55326	Q55326 synechococc
33	20	19.2	15	5	Q9TXC8	Q9txc8 locusta mig
34	20	19.2	15	12	Q84332	Q84332 simian viru
35	20	19.2	16	4	Q9UCI8	Q9uci8 homo sapien
36	20	19.2	16	12	Q84353	Q84353 simian viru
37	20	19.2	17	2	O52748	O52748 anabaena sp
38	20	19.2	18	11	Q9QVB0	Q9qvb0 rattus sp.
39	20	19.2	19	2	Q06140	Q06140 neisseria m
40	20	19.2	19	2	Q9R4X3	Q9r4x3 bacillus ce
41	20	19.2	19	12	Q84863	Q84863 unidentifie
42	20	19.2	20	2	Q9R9A5	Q9r9a5 nitrosospir
43	20	19.2	20	2	Q9R987	Q9r987 nitrosospir
44	20	19.2	20	2	Q9R5T8	Q9r5t8 campylobact
45	20	19.2	20	3	O13594	O13594 saccharomyc

ALIGNMENTS

```
RESULT   1
Q9S8G6
ID   Q9S8G6      PRELIMINARY;      PRT;    19 AA.
AC   Q9S8G6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   TYPE 2 LECTIN (FRAGMENT).
OS   Colocasia esculenta (Elephant's ear) (Taro).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Araceae; Colocasia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95288362.
RA   Van Damme E.J., Goossens K., Smeets K., van Leuven F., Verhaert P.,
RA   Peumans W.J.;
RT   "The major tuber storage protein of araceae species is a lectin.
RT   Characterization and molecular cloning of the lectin from Arum
RT   maculatum L.";
RL   Plant Physiol. 107:1147-1158(1995).
SQ   SEQUENCE   19 AA;  2066 MW;  F9C18865CA58608A CRC64;

  Query Match             27.9%;  Score 29;  DB 10;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKYG 17
           ::  ||| ||
Db       8 LLFSGQVXYG 17


RESULT   2
O03888
ID   O03888      PRELIMINARY;      PRT;    17 AA.
AC   O03888;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   NADH DEHYDROGENASE (EC 1.6.99.3) (CYTOCHROME C REDUCTASE)
DE   (TYPE I DEHYDROGENASE) (FRAGMENT).
GN   NAD5.
```

```
RP   SEQUENCE.
RC   STRAIN=OM5;
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 = CO(2) +
CC       2 H(+) + FERRICYTOCHROME B-561.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0141; PL0141.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1779 MW;  82DD3BF93E739D63 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKY 16
           ||: |   |
Db       1 MMIPGSFDY 9


RESULT  14
DCMM_PSECH
ID   DCMM_PSECH     STANDARD;      PRT;    15 AA.
AC   P19917;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydohydrogena.
OC   Bacteria; Proteobacteria.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0143; PL0143.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1815 MW;  90508CF93E739D63 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKY 16
           ||: |   |
Db       1 MMIPGHFDY 9


RESULT  15
ARCD_PSEPU
ID   ARCD_PSEPU     STANDARD;      PRT;    16 AA.
AC   P41147;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   ARGININE/ORNITHINE ANTIPORTER (FRAGMENT).
GN   ARCD.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4359;
RA   Wilson S.D., Wang M., Filpula D.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE
CC       AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC       ARGININE DEIMINASE PATHWAY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC       PERMEASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07185; AAA16963.1; -.
KW   Transport; Amino-acid transport; Transmembrane; Inner membrane.
FT   NON_TER      1      1
SQ   SEQUENCE   16 AA;  1644 MW;  90B48A7C8FAA9705 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLHN 20
           |||::
Db       7 YGLYD 11


Search completed: December 21, 2000, 08:32:59
Job time: 442 sec
```

```
RESULT  10
TEMA_RANTE
ID   TEMA_RANTE     STANDARD;      PRT;    14 AA.
AC   P56917;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN A.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC       BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1469 MW;  601653612B9DECD4 CRC64;


  Query Match             20.2%;  Score 21;  DB 1;  Length 14;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           |:|:|  |:
Db       4 LIGRVLSGI 12


RESULT  11
TEMF_RANTE
ID   TEMF_RANTE     STANDARD;      PRT;    14 AA.
AC   P56921;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN F.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1441 MW;  4D1653612B9DECC3 CRC64;


  Query Match             20.2%;  Score 21;  DB 1;  Length 14;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           |:|:|  |:
Db       4 LIGKVLSGI 12


RESULT  12
IDHC_PIG
ID   IDHC_PIG       STANDARD;      PRT;    13 AA.
AC   P20304;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   ISOCITRATE DEHYDROGENASE [NADP] CYTOPLASMIC (EC 1.1.1.42)
DE   (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
DE   (FRAGMENT).
GN   IDH1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 87308293.
RA   Bailey J.M., Colman R.F.;
RT   "Isolation of the glutamyl peptide labeled by the nucleotide analogue
RT   2-(4-bromo-2,3-dioxobutylthio)-1,N(6)-ethenoadenosine
RT   2',5'-biphosphate in the active site of NADP+-specific isocitrate
RT   dehydrogenase.";
RL   J. Biol. Chem. 262:12620-12626(1987).
RN   [2]
RP   SEQUENCE.
RX   MEDLINE; 87308292.
RA   Ehrlich R.S., Colman R.F.;
RT   "Characterization of an active site peptide modified by the substrate
RT   analogue 3-bromo-2-ketoglutarate on a single chain of dimeric NADP+-
RT   dependent isocitrate dehydrogenase.";
RL   J. Biol. Chem. 262:12614-12619(1987).
CC   -!- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) = 2-OXOGLUTARATE +
CC       CO(2) + NADPH.
CC   -!- SUBUNIT: HOMODIMER.
DR   PIR; A27371; A27371.
DR   PIR; A27372; A27372.
DR   INTERPRO; IPR001804; -.
DR   PROSITE; PS00470; IDH_IMDH; PARTIAL.
KW   Oxidoreductase; NADP; Tricarboxylic acid cycle.
FT   NON_TER       1       1
FT   ACT_SITE      5       5       POTENTIAL.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1353 MW;  1B640F0E9F7C71E0 CRC64;


  Query Match             19.7%;  Score 20.5;  DB 1;  Length 13;
  Best Local Similarity   45.5%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      10 LLGQVKYGLHN 20
           | |:: :|| |
Db       2 LAGEI-HGLSN 11


RESULT  13
DCMM_PSECA
ID   DCMM_PSECA     STANDARD;      PRT;    15 AA.
AC   P19920;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   CARBON MONOXIDE OXYGENASE (CYTOCHROME B-561) MEDIUM CHAIN (EC 1.2.2.4)
DE   (FRAGMENT).
OS   Pseudomonas carboxydovorans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Oligotropha.
RN   [1]
```

```
CRBL_VESOR
ID   CRBL_VESOR     STANDARD;      PRT;    14 AA.
AC   P17236;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1990 (Rel. 15, Last annotation update)
DE   HISTAMINE RELEASING PEPTIDE II (HR-II).
OS   Vespa orientalis (Oriental hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Miroshnikov A.I., Snezhkova L.G., Nazimov I.V., Reshetova O.I.,
RA   Rozynov B.V., Gushchin I.S.;
RT   "Structure and properties of histamine releasing peptides from the
RT   venom of Vespa orientalis hornet.";
RL   Bioorg. Khim. 7:1467-1477(1981).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
DR   PIR; JN0390; JN0390.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1524 MW;  22015B4A6CEDFD38 CRC64;

  Query Match             21.2%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 9e+02;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 MLLGQVKYGL 18
           ::||::  ||
Db       4 LILGKLVKGL 13


RESULT   7
LECB_IRIHO
ID   LECB_IRIHO     STANDARD;      PRT;    20 AA.
AC   P36231;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   N-ACETYL-D-GALACTOSAMINE-BINDING LECTIN SUBUNIT B (A-DISACCHARIDE-
DE   BINDING LECTIN SUBUNIT B) (FRAGMENT).
OS   Iris hollandica (Dutch iris).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Asparagales; Iridaceae; Iris.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. GOLDEN HARVEST, AND CV. PROF. BLAAUW; TISSUE=BULB;
RX   MEDLINE; 94171801.
RA   Mo H., van Damme E.J.M., Peumans W.J., Goldstein I.J.;
RT   "Isolation and characterization of an
RT   N-acetyl-D-galactosamine-binding lectin from Dutch Iris bulbs which
RT   recognizes the blood group A disaccharide (GalNAc alpha 1-3Gal).";
RL   J. Biol. Chem. 269:7666-7673(1994).
CC   -!- FUNCTION: GAL / GALNAC-SPECIFIC LECTIN. AGGLUTINATES BOTH NATIVE
CC       AND TRYPSIN-TREATED RABBIT ERYTHROCYTES BUT NOT HUMAN ERYTHROCYTES
CC       IRRESPECTIVE OF BLOOD GROUP TYPE.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A AND B CHAINS.
KW   Lectin.
FT   NON_TER       20      20
SQ   SEQUENCE   20 AA;  2246 MW;  DE08DA025FD17D56 CRC64;

  Query Match             21.2%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLG 12
           |: || :|
Db       9 EETMMRIG 16


RESULT   8
AL16_CARMA
ID   AL16_CARMA     STANDARD;      PRT;     8 AA.
AC   P81819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 16.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  813 MW;  7C286B45AB476878 CRC64;

  Query Match             20.2%;  Score 21;  DB 1;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           |   |||
Db       2 GPYSYGL 8


RESULT   9
CRBL_VESTR
ID   CRBL_VESTR     STANDARD;      PRT;    13 AA.
AC   P17231;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VESPID CHEMOTACTIC PEPTIDE T (VESCP-T).
OS   Vespa tropica (Hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Yasuhara T., Nakajima T., Erspaer V.;
RL   (In) Sakakibara S. (eds.);
RL   Peptide chemistry 1982, pp.213-218, Protein Research Foundation,
RL   Osaka (1983).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES       13      13       AMIDATION.
SQ   SEQUENCE   13 AA;  1354 MW;  220140365DFE5338 CRC64;

  Query Match             20.2%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           :||::  ||
Db       4 ILGKILGGL 12
```

```
CC       METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC       3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
CC   -!- PATHWAY: NONOXYDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
DR   PIR; A11497; A11497.
DR   INTERPRO; IPR001585; -.
DR   PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR   PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW   Transferase; Pentose shunt.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1033 MW;  325A31A44EB1E058 CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLH 19
           ||:|
Db       1 YGIH 4


RESULT   3
MDH_KIBAR
ID   MDH_KIBAR      STANDARD;      PRT;    20 AA.
AC   P19978;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Kibdelosporangium aridum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC   Kibdelosporangium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89374824.
RA   Rommel T.O., Hund H.-K., Speth A.R., Lingens F.;
RT   "Purification and N-terminal amino-acid sequences of bacterial malate
RT   dehydrogenases from six actinomycetales strains and from
RT   Phenylobacterium immobile, strain E.";
RL   Biol. Chem. Hoppe-Seyler 370:763-768(1989).
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   PIR; S04961; S04961.
DR   HSSP; P10584; 1BDM.
DR   INTERPRO; IPR001252; -.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  1989 MW;  79587E6F58C00AC9 CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20


RESULT   4
MDH_MICGL
ID   MDH_MICGL      STANDARD;      PRT;    20 AA.
AC   P19979;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Microtetraspora glauca.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptosporangineae; Streptosporangiaceae;
OC   Microtetraspora.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89374824.
RA   Rommel T.O., Hund H.-K., Speth A.R., Lingens F.;
RT   "Purification and N-terminal amino-acid sequences of bacterial malate
RT   dehydrogenases from six actinomycetales strains and from
RT   Phenylobacterium immobile, strain E.";
RL   Biol. Chem. Hoppe-Seyler 370:763-768(1989).
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   PIR; S04958; S04958.
DR   HSSP; P10584; 1BDM.
DR   INTERPRO; IPR001236; -.
DR   INTERPRO; IPR001252; -.
DR   PFAM; PF00056; ldh; 1.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2065 MW;  740129BE59D01EBE CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      11 GQIGYAL 17


RESULT   5
AL11_CARMA
ID   AL11_CARMA     STANDARD;      PRT;     9 AA.
AC   P81814;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 11.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  927 MW;  832D79CDCB46D861 CRC64;


  Query Match             21.2%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||  :||
Db       3 GQYAFGL 9


RESULT   6
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:59 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-13
Perfect score: 104
Sequence: 1 DITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23	22.1	8	1	AL17_CARMA	P81820 carcinus ma
2	23	22.1	9	1	TAL3_PICJA	P17441 pichia jadi
3	23	22.1	20	1	MDH_KIBAR	P19978 kibdelospor
4	23	22.1	20	1	MDH_MICGL	P19979 microtetras
5	22	21.2	9	1	AL11_CARMA	P81814 carcinus ma
6	22	21.2	14	1	CRBL_VESOR	P17236 vespa orien
7	22	21.2	20	1	LECB_IRIHO	P36231 iris hollan
8	21	20.2	8	1	AL16_CARMA	P81819 carcinus ma
9	21	20.2	13	1	CRBL_VESTR	P17231 vespa tropi
10	21	20.2	14	1	TEMA_RANTE	P56917 rana tempor
11	21	20.2	14	1	TEMF_RANTE	P56921 rana tempor
12	20.5	19.7	13	1	IDHC_PIG	P20304 sus scrofa
13	20	19.2	15	1	DCMM_PSECA	P19920 pseudomonas
14	20	19.2	15	1	DCMM_PSECH	P19917 pseudomonas
15	20	19.2	16	1	ARCD_PSEPU	P41147 pseudomonas
16	20	19.2	19	1	PHLC_STAIN	P80924 staphylococ
17	20	19.2	19	1	UP25_UPEIN	P82031 uperoleia i
18	20	19.2	20	1	PGK_CLOPA	P81346 clostridium
19	20	19.2	20	1	UCRQ_EQUAR	P81247 equisetum a
20	19	18.3	10	1	TKS1_AEDAE	P42634 aedes aegyp
21	19	18.3	10	1	TKS2_AEDAE	P42635 aedes aegyp
22	19	18.3	14	1	TEMC_RANTE	P56918 rana tempor
23	19	18.3	14	1	UC18_MAIZE	P80624 zea mays (m
24	19	18.3	15	1	RBS_PHYPA	P80657 physcomitre
25	19	18.3	17	1	BOL4_MEGPE	P07495 megabombus
26	19	18.3	18	1	LUXB_KRYAS	P18300 kryptophana
27	19	18.3	18	1	RL23_HALCU	P05975 halobacteri
28	19	18.3	19	1	MDH_COMAC	P80539 comamonas a
29	19	18.3	20	1	COG2_CHIOP	P34154 chionoecete
30	19	18.3	20	1	COG3_CHIOP	P34155 chionoecete
31	19	18.3	20	1	COGA_PARCM	P20732 paralithode
32	19	18.3	20	1	COGB_PARCM	P20733 paralithode
33	18	17.3	9	1	TAL1_PICJA	P17440 pichia jadi
34	18	17.3	12	1	PA2B_VIPBO	P31859 vipera beru
35	18	17.3	13	1	CRBL_VESMA	P17232 vespa manda
36	18	17.3	14	1	TAT_HV1W2	P12509 human immun
37	18	17.3	14	1	TAT_HV1Z8	P12511 human immun
38	18	17.3	15	1	CDN5_LITCE	P82077 litoria cae
39	18	17.3	17	1	FLA2_BARBA	P35634 bartonella
40	18	17.3	17	1	NU4M_TRIRU	Q36834 trichophyto
41	18	17.3	19	1	RS19_SPICI	O31159 spiroplasma
42	18	17.3	20	1	COXQ_ONCMY	P80335 oncorhynchu
43	18	17.3	20	1	COXQ_THUOB	P80983 thunnus obe
44	17	16.3	8	1	AL15_CARMA	P81818 carcinus ma
45	17	16.3	9	1	DSIP_RABIT	P01158 oryctolagus

ALIGNMENTS

RESULT 1
AL17_CARMA
ID AL17_CARMA STANDARD; PRT; 8 AA.
AC P81820;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CARCINUSTATIN 17.
OS Carcinus maenas (Common shore crab) (Green crab).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN [1]
RP SEQUENCE.
RC TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX MEDLINE; 98121193.
RA Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA Thorpe A.;
RT "Isolation and identification of multiple neuropeptides of the
RT allatostatin superfamily in the shore crab Carcinus maenas.";
RL Eur. J. Biochem. 250:727-734(1997).
CC -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW Neuropeptide; Amidation; Multigene family.
FT MOD_RES 8 8 AMIDATION (POTENTIAL).
SQ SEQUENCE 8 AA; 858 MW; C82879D5AB46D865 CRC64;

Query Match 22.1%; Score 23; DB 1; Length 8;
Best Local Similarity 57.1%; Pred. No. 8.8e+04;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      12 GQVKYGL 18
           ||  :||
Db       2 GQYSFGL 8
```

RESULT 2
TAL3_PICJA
ID TAL3_PICJA STANDARD; PRT; 9 AA.
AC P17441;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).
OS Pichia jadinii (Yeast) (Candida utilis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Pichia.
RN [1]
RP SEQUENCE.
RX MEDLINE; 75145197.
RA Tsolas O., Sun S.C.;
RT "Isolation of a peptide containing a histidinyl-cysteinyl sequence
RT from the active center of transaldolase.";
RL Arch. Biochem. Biophys. 167:525-533(1975).
CC -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF

```
  Matches    4;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMMLLG 12
           |||    :: |
Db       5 ITGNTRKLIFG 15


RESULT  14
A49887
transcription factor HNF-3 beta - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
C;Accession: A49887
R;Sasaki, H.; Hogan, B.L.M.
Cell 76, 103-115, 1994
A;Title: HNF-3beta as a regulator of floor plate development.
A;Reference number: A49887; MUID:94116056
A;Accession: A49887
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-15 <SAS>
C;Keywords: alternative initiators; alternative splicing; transcription factor


  Query Match             21.2%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 AMMLLGQVK 15
           |  :|| ||
Db       4 ASSMLGAVK 12


RESULT  15
S02455
DNA-invertase - phage Mu (fragment)
C;Species: phage Mu
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 16-Feb-1997
C;Accession: S02455
R;Klippel, A.; Mertens, G.; Patschinsky, T.; Kahmann, R.
EMBO J. 7, 1229-1237, 1988
A;Title: The DNA invertase Gin of phage Mu: formation of a covalent complex with DNA via
A;Reference number: S02455; MUID:88296429
A;Accession: S02455
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-16 <KLI>
C;Genetics:
A;Gene: gin
C;Keywords: DNA binding


  Query Match             21.2%;  Score 22;  DB 2;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 MLLGQVK 15
           ||:| |:
Db       1 MLIGYVR 7


Search completed: December 21, 2000, 08:30:07
Job time: 271 sec
```

```
C;Keywords: oxidoreductase; tricarboxylic acid cycle


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      11 GQIGYAL 17


RESULT   9
B44835
dTDPglucose 4,6-dehydratase (EC 4.2.1.46) - Streptomyces peucetius (fragment)
C;Species: Streptomyces peucetius
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: B44835
R;Thompson, M.W.; Strohl, W.R.; Floss, H.G.
J. Gen. Microbiol. 138, 779-786, 1992
A;Title: Purification and characterization of TDP-D-glucose 4,6-dehydratase from anthrac
A;Reference number: A44835; MUID:92268857
A;Accession: B44835
A;Molecule type: protein
A;Residues: 1-20 <THO>
A;Cross-references: PIDN:AAB22196.1; PID:g249564
A;Experimental source: ATCC 29050
A;Note: sequence extracted from NCBI backbone (NCBIP:103673)
C;Keywords: carbon-oxygen lyase; hydro-lyase


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20


RESULT  10
PH1338
Ig heavy chain DJ region (clone C372-115) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1338
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor lymph
A;Reference number: PH1302; MUID:93094761
A;Accession: PH1338
A;Molecule type: DNA
A;Residues: 1-20 <WAS>
C;Keywords: heterotetramer; immunoglobulin


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   30.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKYG 17
           :: :| ::||
Db       4 ILRIGLLRYG 13


RESULT  11
JN0390
histamine-releasing peptide II - oriental hornet
N;Alternate names: venom protein HR-2
C;Species: Vespa orientalis (oriental hornet)
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 23-Aug-1997
C;Accession: JN0390; S10919
R;Miroshnikov, A.I.; Snezhkova, L.G.; Nazimov, I.V.; Reshetova, O.I.; Rozinov, B.V.; Gus
Bioorg. Khim. 7, 1467-1477, 1981
A;Title: Structure and properties of histamine releasing peptides from the venom of V
A;Reference number: JN0389
A;Accession: JN0390
A;Molecule type: protein
A;Residues: 1-14 <MIR>
R;Tuichibaev, M.U.; Akhmedova, N.U.; Kazakov, I.; Korneev, A.S.; Gagel'gans, A.I.
Biochemistry (N.Y.) 53, 183-190, 1988
A;Title: Low-molecular-weight peptides of venom of the giant hornet Vespa orientalis.
A;Reference number: S06445
A;Accession: S10919
A;Molecule type: protein
A;Residues: 1-14 <TUI>
C;Superfamily: crabrolin
C;Keywords: amidated carboxyl end; venom
F;14/Modified site: amidated carboxyl end (Leu) #status experimental


  Query Match             21.2%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 MLLGQVKYGL 18
           ::||::  ||
Db       4 LILGKLVKGL 13


RESULT  12
S72217
D-arabinose 1-dehydrogenase (NAD(P)+) (EC 1.1.1.117) - yeast (Candida albicans) (frag
C;Species: Candida albicans
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 13-Mar-1998
C;Accession: S72217
R;Kim, S.T.; Huh, W.K.; Kim, J.Y.; Hwang, S.W.; Kang, S.O.
Biochim. Biophys. Acta 1297, 1-8, 1996
A;Title: D-arabinose dehydrogenase and biosynthesis of erythroascorbic acid in Candid
A;Reference number: S72217; MUID:96439039
A;Accession: S72217
A;Molecule type: protein
A;Residues: 1-14 <KIM>
C;Keywords: oxidoreductase


  Query Match             21.2%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   30.8%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKYGLHN 20
           | |  :: : |:|
Db       1 MKLATEIDFXLNN 13


RESULT  13
PH0797
T-cell receptor alpha chain (PF2.10.1 V-alpha-3.AR5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0797
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility comple
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0797
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60903
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor


  Query Match             21.2%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   36.4%;  Pred. No. 2e+03;
```

```
RESULT   3
PC4213
bphB protein - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 17-Dec-1996 #sequence_revision 21-Jan-1997 #text_change 08-Oct-1999
C;Accession: PC4213
R;Sylvestre, M.; Sirois, M.; Hurtubise, Y.; Bergeron, J.; Ahmad, D.; Shareck, F.; Barria
Gene 174, 195-202, 1996
A;Title: Sequencing of Comamonas testosteroni strain B-356-biphenyl/chlorobiphenyl dioxy
A;Reference number: JC4993; MUID:97045812
A;Accession: PC4213
A;Molecule type: DNA
A;Residues: 1-15 <SYL>
A;Cross-references: GB:U47637; NID:g1245151; PIDN:AAC44530.1; PID:g1245156
A;Experimental source: strain B-356
C;Genetics:
A;Gene: bphB


  Query Match             23.1%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 9.1e+02;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMMLLG 12
           :||| |::  |
Db       3 LTGEVALITGG 13


RESULT   4
PH1475
T-cell receptor beta chain (clone 223/5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1475
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kc
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1475
A;Molecule type: mRNA
A;Residues: 1-16 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             23.1%;  Score 24;  DB 2;  Length 16;
  Best Local Similarity   40.0%;  Pred. No. 9.8e+02;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLG 12
           |||:  :  |
Db       7 TGERERLFFG 16


RESULT   5
A11497
transaldolase (EC 2.2.1.2) III - yeast (Pichia jadinii) (fragment)
C;Species: Pichia jadinii, Candida utilis
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
C;Accession: A11497
R;Tsolas, O.; Sun, S.C.
Arch. Biochem. Biophys. 167, 525-533, 1975
A;Title: Isolation of a peptide containing a histidinyl-cysteinyl sequence from the acti
A;Reference number: A11497; MUID:75145197
A;Accession: A11497
A;Molecule type: protein
A;Residues: 1-9 <TSO>
C;Keywords: transferase


  Query Match             22.1%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLH 19
           ||:|
Db       1 YGIH 4


RESULT   6
S74147
glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S74147
R;Fukuda, A.; Osawa, T.; Hitomi, K.; Uchida, K.
Arch. Biochem. Biophys. 333, 419-426, 1996
A;Title: 4-Hydroxy-2-nonenal cytotoxicity in renal proximal tubular cells: protein mo
A;Reference number: S74147; MUID:96404942
A;Accession: S74147
A;Molecule type: protein
A;Residues: 1-10 <FUK>
A;Experimental source: LLC-PK1 cells (renal tubular epithelial cells)
C;Keywords: NAD; oxidoreductase


  Query Match             22.1%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 8.7e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GQVKYGLH 19
           |:|| |::
Db       1 GKVKVGVN 8


RESULT   7
S04961
malate dehydrogenase (EC 1.1.1.37) - Kibdelosporangium aridum (fragment)
C;Species: Kibdelosporangium aridum
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Aug-1994
C;Accession: S04961
R;Rommel, T.O.; Hund, H.K.; Speth, A.R.; Lingens, F.
Biol. Chem. Hoppe-Seyler 370, 763-768, 1989
A;Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydro
A;Reference number: S04956; MUID:89374824
A;Accession: S04961
A;Molecule type: protein
A;Residues: 1-20 <ROM>
C;Superfamily: L-lactate dehydrogenase
C;Keywords: oxidoreductase; tricarboxylic acid cycle


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20


RESULT   8
S04958
malate dehydrogenase (EC 1.1.1.37) - Microtetraspora glauca (fragment)
C;Species: Microtetraspora glauca
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Aug-1994
C;Accession: S04958
R;Rommel, T.O.; Hund, H.K.; Speth, A.R.; Lingens, F.
Biol. Chem. Hoppe-Seyler 370, 763-768, 1989
A;Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydro
A;Reference number: S04956; MUID:89374824
A;Accession: S04958
A;Molecule type: protein
A;Residues: 1-20 <ROM>
C;Superfamily: L-lactate dehydrogenase
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:07 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-13
Perfect score: 104
Sequence: 1 DITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	25.0	18	2	I52623	hypothetical prote
2	24	23.1	15	2	PA0091	methionine adenosy
3	24	23.1	15	2	PC4213	bphB protein - Com
4	24	23.1	16	2	PH1475	T-cell receptor be
5	23	22.1	9	2	A11497	transaldolase (EC
6	23	22.1	10	2	S74147	glyceraldehyde-3-p
7	23	22.1	20	2	S04961	malate dehydrogena
8	23	22.1	20	2	S04958	malate dehydrogena
9	23	22.1	20	2	B44835	dTDPglucose 4,6-de
10	23	22.1	20	2	PH1338	Ig heavy chain DJ
11	22	21.2	14	2	JN0390	histamine-releasin
12	22	21.2	14	2	S72217	D-arabinose 1-dehy
13	22	21.2	15	2	PH0797	T-cell receptor al
14	22	21.2	15	2	A49887	transcription fact
15	22	21.2	16	2	S02455	DNA-invertase - ph
16	22	21.2	16	2	PH0773	T-cell receptor be
17	22	21.2	16	2	PH1476	T-cell receptor be
18	22	21.2	20	2	A47105	dystroglycan - chi
19	22	21.2	20	2	S10876	hypothetical prote
20	21	20.2	10	2	PQ0785	NADH dehydrogenase
21	21	20.2	11	2	A58502	38K kidney stone p
22	21	20.2	16	2	D45193	zinc finger protei
23	21	20.2	18	2	A41877	LcrKc - Yersinia p
24	21	20.2	18	2	A49857	megakaryocyte pote
25	21	20.2	19	2	B61056	cytochrome P450 PB
26	21	20.2	20	2	PH1358	Ig heavy chain DJ
27	21	20.2	20	2	A60801	acrosome stabilizi
28	20	19.2	9	2	G58502	kidney and bladder
29	20	19.2	12	2	PH1635	Ig H chain V-D-J r
30	20	19.2	13	2	PS0325	tetrahydroberberin
31	20	19.2	13	2	PH1772	T cell receptor al
32	20	19.2	14	2	PS0371	hypothetical prote
33	20	19.2	14	2	PH1332	Ig heavy chain DJ
34	20	19.2	15	2	S26791	Ig heavy chain V r
35	20	19.2	15	2	PL0143	carbon-monoxide de
36	20	19.2	15	2	A30330	neuropeptide pep -
37	20	19.2	15	2	PH1619	Ig H chain V-D-J r
38	20	19.2	15	2	PH0775	T-cell receptor al
39	20	19.2	15	2	PH0779	T-cell receptor al
40	20	19.2	15	2	PH1455	T-cell receptor al
41	20	19.2	18	2	A39997	group III allergen
42	20	19.2	19	2	S43657	hsp90 protein homo
43	20	19.2	20	2	S11416	ribosomal protein
44	19.5	18.8	15	2	PA0106	protein QF200076 -
45	19.5	18.8	19	2	A39504	octamer-binding pr

ALIGNMENTS

RESULT 1
I52623
hypothetical protein TCR delta [imported] - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000
C;Accession: I52623
R;Przybylski, G.; Oettle, H.; Ludwig, W.D.; Siegert, W.; Schmidt, C.A.
Br. J. Haematol. 87, 301-307, 1994
A;Title: Molecular characterization of illegitimate TCR delta gene rearrangements in
A;Reference number: I52623; MUID:95034253
A;Accession: I52623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <RES>
A;Cross-references: GB:S73537; NID:g688155
C;Genetics:
A;Gene: TCRdelta

Query Match 25.0%; Score 26; DB 2; Length 18;
Best Local Similarity 41.2%; Pred. No. 5.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

```
Qy      11 LGQVKY--------GLH 19
           ||| ||         |:|
Db       1 LGQKKYSSYLWGTSGVH 17
```

RESULT 2
PA0091
methionine adenosyltransferase (EC 2.5.1.6) 2 - fungus (Fusarium sporotrichioides) (f
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-May-2000
C;Accession: PA0091
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Reference number: PA0051
A;Accession: PA0091
A;Molecule type: protein
A;Residues: 1-15 <CHO>
C;Keywords: S-adenosylmethionine; transferase

Query Match 23.1%; Score 24; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 9.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 TGEKAMMLLG 12
           | |||:  ||
Db       1 TYEKALEQLG 10
```

CC uraemia, diabetes, Cushing's syndrome, cachexias, eating disorders, AIDS,
CC after stress and during neuromuscular disease.
XX
SQ Sequence 12 AA;

Query Match 27.9%; Score 31; DB 20; Length 12;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy      11 EHLLVDFLQS 20
           | |||:||::
Db       1 edllveflen 10
```

Search completed: December 21, 2000, 08:36:13
Job time: 472 sec

```
CC   apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,
CC   which forms an amphipathic alpha-helix in presence of lipids. (A),
CC   optionally as a complex with lipids, and host cells that contain (A),
CC   are useful for gene therapy, or prevention, of diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I
CC   deficiency, hypertriglyceridemia and metabolic syndrome. also to treat
CC   endotoxemia (septic shock). Host cells containing (A) can also be used
CC   to study the role of apoA-I in lipid metabolism. (B) can be used
CC   diagnostically, e.g. to measure serum HDL (particularly its
CC   subpopulation involved in retrograde cholesterol transport) and for
CC   imaging the circulatory system or HDL accumulations at fatty streaks.
CC   The present sequence represents a peptide from the present invention.
XX
SQ   Sequence   22 AA;


  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  14
Y19320
ID   Y19320 standard; Peptide; 22 AA.
XX
AC   Y19320;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW   Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW   human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW   cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW   high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW   septic shock.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916459-A1.
XX
PD   08-APR-1999.
XX
PF   28-SEP-1998;   98WO-US20327.
XX
PR   29-SEP-1997;   97US-0940095.
XX
PA   (BUTT/) BUTTNER K.
PA   (CORN/) CORNUT I.
PA   (DASS/) DASSEUX J.
PA   (DUFO/) DUFOURCQ J.
PA   (METZ/) METZ G.
PA   (SEKU/) SEKUL R.
XX
PI   Buttner K,  Cornut I,  Dasseux J,  Dufourcq J,  Metz G;
PI   Sekul R;
XX
DR   WPI; 1999-277035/23.
XX
PT   Peptide agonists of apolipoprotein A-I
XX
PS   Claim 19; Page 157; 280pp; English.
XX
CC   The present invention describes an agonist (A) of apolipoprotein A-I
CC   (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC   amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC   complexes, are used to treat or prevent diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC   deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC   treating septic shock. When labeled, (A) can also be used diagnostically
CC   to measure serum levels of HDL, in particular the HDL subpopulation that
CC   is involved in retrograde cholesterol transport, also to image HDL at
CC   e.g. atherosclerotic streaks, and to raise antibodies. Y19188 to Y19441
CC   represent lecithin:cholesterol acyltransferase (LCAT) activity
CC   exhibiting core peptides, which are apoA-I agonists.
XX
SQ   Sequence   22 AA;


  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  15
Y32796
ID   Y32796 standard; peptide; 12 AA.
XX
AC   Y32796;
XX
DT   09-NOV-1999  (first entry)
XX
DE   Mammalian prolactin receptor derived polypeptide.
XX
KW   Signal transduction; proteolytic cleavage; prolactin receptor;
KW   proteosome binding site; muscle wasting; renal tubular defect; uraemia;
KW   diabetes; Cushing's disease; eating disorder; AIDS; ubiquitin;
KW   growth hormone deficiency; inhibitor.
XX
OS   Oryctolagus cuniculus.
OS   Rattus sp.
OS   Mus sp.
XX
PN   EP943624-A1.
XX
PD   22-SEP-1999.
XX
PF   12-MAR-1998;   98EP-0200799.
XX
PR   12-MAR-1998;   98EP-0200799.
XX
PA   (UYUT-) RIJKSUNIV UTRECHT.
XX
DR   WPI; 1999-510568/43.
XX
PT   Controlling the availability and/or signal transduction capability
PT   of a cell surface receptor, useful for treating growth hormone
PT   deficiencies
XX
PS   Disclosure; Page 9; 36pp; English.
XX
CC   This sequence is a polypeptide from the mammalian prolactin receptor.
CC   Sequences (Y32794-Y32823) are examples of polypeptide sequences found at
CC   or near the ubiquitin/proteosome complex binding site located on the
CC   intracellular part of a cell surface receptor. These sequences are used
CC   in a method for controlling the availability and signal transduction
CC   capability of a cell surface receptor by administering an inhibitor that
CC   is capable of inhibiting proteolytic cleavage of the receptor. Inhibition
CC   of this proteolytic cleavage results in the receptors being present on
CC   the surface for longer and therefore signalling for longer to the
CC   interior of the cell. This increases the sensitivity of cells to any
CC   hormones which might be present. The inhibitor is either derived from,
CC   competes with or binds to a polypeptide sequence of which sequences
CC   Y32794-Y32823 are examples. The inhibitor may be used to treat muscle
CC   wasting, associated with disorders such as renal tubular defects,
```

PA (METZ/) METZ G.
PA (SEKU/) SEKUL R.
XX
PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;
XX
DR WPI; 1999-277034/23.
XX
PT Peptide agonists of apolipoprotein A-I
XX
PS Example; Page 114; 254pp; English.
XX
CC The present invention describes an agonist (A) of apolipoprotein A-I
CC (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC complexes, are used to treat or prevent diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC treating septic shock. When labeled, (A) can also be used diagnostically
CC to measure serum levels of HDL, in particular the HDL subpopulation that
CC is involved in retrograde cholesterol transport, also to image HDL at
CC e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187
CC represent lecithin:cholesterol acyltransferase (LCAT) activity
CC exhibiting core peptides, which are apoA-I agonists.
XX
SQ Sequence 22 AA;

Query Match 28.8%; Score 32; DB 20; Length 22;
Best Local Similarity 58.3%; Pred. No. 94;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 11 EHLLVDFLQSLS 22
| || | ||:|:
Db 7 erlledllqaln 18

RESULT 12
Y18812
ID Y18812 standard; Peptide; 22 AA.
XX
AC Y18812;
XX
DT 09-JUL-1999 (first entry)
XX
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW septic shock.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916408-A2.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-US20328.
XX
PR 29-SEP-1997; 97US-0940093.
XX
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (DASS/) DASSEUX J.
PA (METZ/) METZ G.
PA (SEKU/) SEKUL R.
XX
PI Buttner K, Cornut I, Dasseux J, Metz G, Sekul R;
XX
DR WPI; 1999-277031/23.
XX
PT Peptide agonists of apolipoprotein A-I
XX
PS Example; Page 111; 152pp; English.
XX
CC The present invention describes an agonist (A) of apolipoprotein A-I
CC (apoA-I) which is a 14-22 residue peptide, or analog, that forms an
CC amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC complexes, are used to treat or prevent diseases associated with
CC dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC treating septic shock. When labeled, (A) can also be used diagnostically
CC to measure serum levels of HDL, in particular the HDL subpopulation that
CC is involved in retrograde cholesterol transport, also to image HDL at
CC e.g. atherosclerotic streaks, and to raise antibodies. Y18680 to Y18933
CC represent lecithin:cholesterol acyltransferase (LCAT) activity
CC exhibiting core peptides, which are apoA-I agonists.
XX
SQ Sequence 22 AA;

Query Match 28.8%; Score 32; DB 20; Length 22;
Best Local Similarity 58.3%; Pred. No. 94;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 11 EHLLVDFLQSLS 22
| || | ||:|:
Db 7 erlledllqaln 18

RESULT 13
Y18549
ID Y18549 standard; Peptide; 22 AA.
XX
AC Y18549;
XX
DT 09-JUL-1999 (first entry)
XX
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;
KW ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;
KW hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916409-A2.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-US20329.
XX
PR 29-SEP-1997; 97US-0940136.
XX
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (DASS/) DASSEUX J.
PA (DUFO/) DUFOURCQ J.
PA (METZ/) METZ G.
PA (SEKU/) SEKUL R.
XX
PI Buttner K, Cornut I, Dasseux J, Dufourcq J, Metz G;
PI Sekul R;
XX
DR WPI; 1999-254921/21.
XX
PT Nucleic acid encoding apolipoprotein A-I agonist peptides
XX
PS Claim 18; Page 166; 232pp; English.
XX
CC The present invention describes a nucleic acid (A) encoding an

PR 17-NOV-1997; 97US-0066090.
PR 17-NOV-1997; 97US-0066094.
PR 17-NOV-1997; 97US-0066095.
PR 17-NOV-1997; 97US-0066089.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;
PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y;
XX
DR WPI; 1999-337740/28.
DR N-PSDB; X85016.
XX
PT New human secreted proteins and coding sequences useful for treating
PT disorders of the immune system and hyperproliferative disorders
XX
PS Disclosure; Page 119; 507pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. X84924) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 125 novel genes and their fragments (nucleic
CC acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 125 polynucleotides, based on
CC which tissues they are most highly expressed in (see X84933 for described
CC uses).
XX
SQ Sequence 14 AA;

Query Match 29.7%; Score 33; DB 20; Length 14;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 10 PEHLLVDFLQSLS 22
| | |||||
Db 1 phssrvsflqsls 13

RESULT 10
W65690
ID W65690 standard; peptide; 21 AA.
XX
AC W65690;
XX
DT 16-OCT-1998 (first entry)
XX
DE Fibronectin binding protein-derived peptide #34.
XX
KW microbial surface components recognising adhesive matrix molecule;
KW MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
KW antibiotic; bacterial infection; antibiotic-resistant strain.
XX
OS Synthetic.
OS Staphylococcus aureus.
XX
PN WO9831389-A2.
XX
PD 23-JUL-1998.
XX
PF 21-JAN-1998; 98WO-US01222.
XX
PR 21-JAN-1997; 97US-0036139.
XX
PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX
PI Hoeoek M, House-pompeo KL, Joh D, Mcgavin MJ, Patti JM;
PI Speziale P;
XX
DR WPI; 1998-413816/35.
XX
PT Antibody that binds to fibronectin-binding protein, preventing its
PT binding to fibronectin - used to treat or prevent bacterial
PT infection, especially by Staphylococci and Streptococci
XX
PS Example 8; Page 101; 201pp; English.
XX
CC The invention relates to antibodies that bind to a fibronectin-binding
CC domain of a fibronectin-binding protein, and inhibit binding of the
CC protein to fibronectin. Also claimed are: (1) isolated peptides of a
CC fibronectin-binding protein that do not bind to fibronectin; (2) fusion
CC protein containing at least one peptide of a fibronectin-binding protein
CC linked to a second amino acid sequence; (3) nucleic acid encoding
CC the isolated peptide of (1). Antibodies, the isolated peptides of
CC (1) and the nucleic acids are all useful for immunisation (active or
CC passive) and (by inhibiting binding of bacteria to fibronectin) for
CC preventing or treating infection in humans or other animals, particularly
CC by staphylococci or streptococci, e.g. meningitis, otitis media,
CC pneumonia, endocarditis, mastitis in cattle, abortion in horses and many
CC others. Since the antibodies block binding of bacteria, they should be
CC effective against antibiotic-resistant strains, and may replace
CC antibiotic therapy or increase its effectiveness. Sequences W65670-90
CC represent a series of synthetic peptides based on the D3 repeat of
CC S. aureus fibronectin binding protein A. They were synthesised to contain
CC a proline residue at each position throught the sequence (ie a proline
CC scan).
XX
SQ Sequence 21 AA;

Query Match 28.8%; Score 32; DB 19; Length 21;
Best Local Similarity 54.5%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 7 FGFPEHLLVDF 17
: || | |||
Db 4 yqfpphnsvdf 14

RESULT 11
Y19066
ID Y19066 standard; Peptide; 22 AA.
XX
AC Y19066;
XX
DT 09-JUL-1999 (first entry)
XX
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW septic shock.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916458-A1.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-US20326.
XX
PR 29-SEP-1997; 97US-0940096.
XX
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (DASS/) DASSEUX J.

PI Mark DF, Creasey AA;
XX
DR WPI; 1983-723186/30.
DR N-PSDB; N30158.
XX
PT Multi:class hybrid interferon poly:peptide(s) - with restricted
PT antiviral and cell growth regulatory activities
XX
PS Example; Fig 17; 61pp; English.
XX
CC The inventors claim a multiclass hybrid interferon polypeptide and a
CC DNA unit having a nucleotide sequence which encodes it. Pref. the
CC AA sequence consists of alpha and beta interferons. Pref. IF1 is
CC (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
CC of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
CC HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
CC N30160, P30227). Alternativeley IF1 is the amino terminal end of a
CC beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC 1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC resp.) (see N30156, P30223). In the examples plasmids pGW5 and
CC pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,
CC N30157). HinfI was used to digest the DNA sequences in the region
CC of significant handicaps (see N30153, N30154, N30158, N30159), and
CC the restriction fragments were ligated to form hybrid DNA.
XX
SQ Sequence 12 AA;

Query Match 29.7%; Score 33; DB 4; Length 12;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DFGFPE 11
|||||:
Db 2 dfgfpq 7

RESULT 8
P80053
ID P80053 standard; protein; 12 AA.
XX
AC P80053;
XX
DT 17-NOV-1990 (first entry)
XX
DE Sequence of human interferon (huIFN) alpha-61A gene around AA 40.
XX
KW Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW antitumour; antiviral; therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 7
FT /note="Residue 40"
XX
PN US4758428-A.
XX
PD 19-JUL-1988.
XX
PF 15-JUL-1985; 85US-0755265.
XX
PR 15-JUL-1985; 85US-0755265.
PR 19-JAN-1983; 83CA-0419758.
XX
PA (CETU) CETUS CORP.
XX
PI Mark DF, Creasey AA;
XX
DR WPI; 1988-219882/31.
DR N-PSDB; n80050.
XX
PT Multi-class hybrid interferon polypeptide(s) -
PT having sequence from interferon-alpha-1 and sequence from
PT interferon-beta-1 for restricted activity
XX
PS Example; Fig 17; 24pp; English.
XX
CC Multi-class hybrid IFN polypeptides having an AA sequence composed
CC of 2 distinct subsequences are claimed. The plasmids used in the
CC construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC pDM101/trp/beta-1. Assembly of the palpha61A plasmid invovled replacing
CC the DNA fragment encoding the 23 AA signal polypeptide of preinterferon
CC with a 120BP EcoRI/Sau3A promoter fragment E.coli trp promoter, operator,
CC and trp leader ribosome binding site preceoperator, encoding an ATG
CC initiation codon and using HindIII site that was inserted, 59 nucleotides
CC 3'-end of the TGA translational stop codon, to insert the gene into the
CC plasmid pBW11 (a deriv of pBR322 having a deletion between the HindIII
CC and PvuII sites). The complete DNA sequence of the promoter and gene
CC fragments inserted between the EcoRI and HindIII sites of pBW11 is shown
CC in n80049. The hybrid gene was constructed by taking advantage of the
CC homologies between huIFN alpha-61A & huIFN beta-1 at around AA 40 of both
CC proteins. The DNA sequence 5'-proximal to the DdeI restriction enzyme
CC cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC 3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC two genes while preserving the translational reading frame of both.
XX
SQ Sequence 12 AA;

Query Match 29.7%; Score 33; DB 9; Length 12;
Best Local Similarity 83.3%; Pred. No. 34;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 6 DFGFPE 11
|||||:
Db 2 dfgfpq 7

RESULT 9
Y27814
ID Y27814 standard; Protein; 14 AA.
XX
AC Y27814;
XX
DT 30-JUL-1999 (first entry)
XX
DE Human secreted protein encoded by gene No. 84.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9924836-A1.
XX
PD 20-MAY-1999.
XX
PF 04-NOV-1998; 98WO-US23435.
XX
PR 17-NOV-1997; 97US-0066100.
PR 07-NOV-1997; 97US-0064900.
PR 07-NOV-1997; 97US-0064908.
PR 07-NOV-1997; 97US-0064911.
PR 07-NOV-1997; 97US-0064912.
PR 07-NOV-1997; 97US-0064983.
PR 07-NOV-1997; 97US-0064984.
PR 07-NOV-1997; 97US-0064985.
PR 07-NOV-1997; 97US-0064987.
PR 07-NOV-1997; 97US-0064988.

```
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
XX
SQ   Sequence   16 AA;


  Query Match             71.2%;  Score 79;  DB 21;  Length 16;
  Best Local Similarity   93.8%;  Pred. No. 3.5e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFLQSLS 22
           ||||:|||||||||||
Db       1 fgfpkhllvdflqsls 16


RESULT   6
W24294
ID   W24294 standard; peptide; 11 AA.
XX
AC   W24294;
XX
DT   17-OCT-1997  (first entry)
XX
DE   Human/Rabbit CETP common peptide.
XX
KW   Immune response; high density lipoprotein; HDL; cholesterol; human;
KW   serum; epitope; cholesteryl ester transfer protein; CETP; rabbit.
XX
OS   Oryctolagus cuniculus.
OS   Homo sapiens.
XX
PN   WO9639168-A1.
XX
PD   12-DEC-1996.
XX
PF   05-JUN-1996;   96WO-US09143.
XX
PR   06-JUN-1995;   95US-0482454.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW,  Carlo DJ,  Kwoh DY;
XX
DR   WPI; 1997-042849/04.
XX
PT   Stimulating an immune response to increase high density lipoprotein
PT   - avoids repeated administration of toxic drugs to lower cholesteryl
PT   ester transfer protein levels
XX
PS   Claim 5; Page 16; 26pp; English.
XX
CC   The sequences given in W24292-94 were used in the method of the
CC   invention to stimulate an immune response to increase high density
CC   lipoprotein (HDL) cholesterol in a mammal exhibiting low levels of
CC   serum HDL.  These peptides represent immunogenic epitopes of
CC   cholesteryl ester transfer protein (CETP).  The method utilises
CC   the body's own immune system to lower CETP levels, thereby increasing
CC   the level of beneficial HDL cholesterol, preferably in serum.  The
CC   method avoids the problems associated with the repeated administration
CC   of drugs which have undesirable side effects.  This peptide represents
CC   a region of CETP which is common to both human and rabbit proteins.
XX
SQ   Sequence   11 AA;


  Query Match             47.7%;  Score 53;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 HLLVDFLQSLS 22
           ||||||||||||
Db       1 hllvdflqsls 11


RESULT   7
P30225
ID   P30225 standard; Protein; 12 AA.
XX
AC   P30225;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -alpha-61A around amino acid 40.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /label= AA No. 40
XX
PN   WO8302461-A.
XX
PD   21-JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU-) CETUS CORP.
XX
```

```
Y91229
ID   Y91229 standard; peptide; 16 AA.
XX
AC   Y91229;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:107.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13975.
XX
PR   20-JUN-1998;   98US-0100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus  -
XX
PS   Claim 10; Page 50; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
XX
SQ   Sequence   16 AA;


  Query Match             74.8%;  Score 83;  DB 21;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFLQSLS 22
           ||||||||||||||||
Db       1 fgfpehllvdflqsls 16


RESULT   5
Y91230
ID   Y91230 standard; peptide; 16 AA.
XX
AC   Y91230;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:108.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13975.
XX
PR   20-JUN-1998;   98US-0100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus  -
XX
PS   Claim 10; Page 62; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
```

```
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             100.0%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||||||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   2
Y13821
ID   Y13821 standard; peptide; 22 AA.
XX
AC   Y13821;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Human CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Homo sapiens.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 88; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the human
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             100.0%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||||||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   3
Y13809
ID   Y13809 standard; peptide; 22 AA.
XX
AC   Y13809;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Example 1; Page 73; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             96.4%;  Score 107;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 2.3e-10;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||:||||||||||
Db      1 lllqmdfgfpkhllvdflqsls 22


RESULT   4
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:12 ; Search time 34.96 Seconds
(without alignments)
21.518 Million cell updates/sec

Title: US-08-934-367-10
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 134459

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	111	100.0	22	20	Y13815	Rabbit CETP immuno
2	111	100.0	22	20	Y13821	Human CETP immunog
3	107	96.4	22	20	Y13809	Rabbit CETP immuno
4	83	74.8	16	21	Y91229	Human cholesteryl
5	79	71.2	16	21	Y91230	Human cholesteryl
6	53	47.7	11	18	W24294	Human/Rabbit CETP
7	33	29.7	12	4	P30225	Sequence of interf
8	33	29.7	12	9	P80053	Sequence of human
9	33	29.7	14	20	Y27814	Human secreted pro
10	32	28.8	21	19	W65690	Fibronectin bindin
11	32	28.8	22	20	Y19066	Lecithin:cholester
12	32	28.8	22	20	Y18812	Lecithin:cholester
13	32	28.8	22	20	Y18549	Lecithin:cholester
14	32	28.8	22	20	Y19320	Lecithin:cholester
15	31	27.9	12	20	Y32796	Mammalian prolacti
16	31	27.9	21	19	W65666	Fibronectin bindin
17	31	27.9	22	19	W65669	Peptide #13. Synt
18	30	27.0	20	15	R61276	Transactivating pr
19	30	27.0	20	19	W41190	Tax protein fragme
20	30	27.0	22	17	R95899	Fragment #4 of 7-a
21	29	26.1	11	20	W99442	Interleukin-2 rece
22	29	26.1	11	21	Y82915	Peptide exhibiting
23	29	26.1	13	18	W10883	MAb anti-HBsAg bin
24	29	26.1	19	16	W13199	Fragment of p53 bi
25	29	26.1	19	19	W57353	Human WBP1 immunog
26	29	26.1	20	16	R72269	Glutamic acid deca
27	29	26.1	20	21	Y59570	GAD65 fragment, pe
28	29	26.1	21	21	Y69793	Human interleukin-
29	28	25.2	9	19	W72564	Dengue virus type-
30	28	25.2	10	15	Y38102	Hepatitis B virus-
31	28	25.2	10	20	Y45670	Immunogenic peptid
32	28	25.2	13	19	W62713	Streptococcus pneu
33	28	25.2	14	13	R21617	Sequence encoded b
34	28	25.2	15	19	W45817	Peptide recognised
35	28	25.2	15	19	W45612	Peptide recognised
36	28	25.2	21	18	W38080	PPPPY motif contai
37	28	25.2	22	20	Y19067	Lecithin:cholester
38	28	25.2	22	20	Y18813	Lecithin:cholester
39	28	25.2	22	20	Y18550	Lecithin:cholester
40	28	25.2	22	20	Y19321	Lecithin:cholester
41	28	25.2	22	20	W96827	Nucleic acid bindi
42	27.5	24.8	16	17	R98414	Interferon-gamma r
43	27	24.3	9	17	R88341	Adenomatous polypo
44	27	24.3	12	16	R78515	Synthetic HTLV pep
45	27	24.3	12	18	W28864	HTLV-1 derived pep

ALIGNMENTS

RESULT 1
Y13815
ID Y13815 standard; peptide; 22 AA.
XX
AC Y13815;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Disclosure; Page 75; 99pp; English.
XX

```
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLGQVK 15
           |||  || ::|
Db       2 EKAKKLLEELK 12


RESULT  15
5304631-16
;Patent No. 5304631
;    APPLICANT: STEWART, JOHN M.;HAHN, KARL W.;. KLIS, WIESLAW A.
;    TITLE OF INVENTION: SYNTHETIC HELIZYME ENZYMES
;    NUMBER OF SEQUENCES: 16
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/464,932
;      FILING DATE: 16-JAN-1990
;SEQ ID NO:16:
;      LENGTH: 15
5304631-16


  Query Match             25.0%;  Score 26;  DB 5;  Length 15;
  Best Local Similarity   54.5%;  Pred. No. 1.6e+02;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLGQVK 15
           |||  || ::|
Db       2 EKAKKLLEELK 12


Search completed: December 21, 2000, 08:31:50
Job time: 373 sec
```

```
  Query Match           25.0%;  Score 26;  DB 4;  Length 13;
  Best Local Similarity  100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 DITGE 5
          |||||
Db      7 DITGE 11


RESULT  12
US-08-484-905-28
; Sequence 28, Application US/08484905
; Patent No. 5976551
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
;    APPLICANT:  Kourilsky, Philippe
;    TITLE OF INVENTION:  An Altered Major Histocompatibility
;    TITLE OF INVENTION:  Complex(MHC) Determinant and Methods for Using the
;    TITLE OF INVENTION:  Determinant
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS-/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,905
;      FILING DATE:  07-JUNE-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Potter, Jane E. R.
;      REGISTRATION NUMBER:  33,332
;      REFERENCE/DOCKET NUMBER:  03495.0106-03000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-905-28


  Query Match           25.0%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity  38.5%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 TGEKAMMLLGQVK 15
          | |    :|||::
Db      3 TSEGCRQILGQLQ 15


RESULT  13
US-08-481-985B-28
; Sequence 28, Application US/08481985B
; Patent No. 6011146
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
;    APPLICANT:  Kourilsky, Phillipe
;    TITLE OF INVENTION:  Altered Major Histocompatibility Complex
;    TITLE OF INVENTION:
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,985B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  03495.0106-04000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-481-985B-28


  Query Match           25.0%;  Score 26;  DB 3;  Length 15;
  Best Local Similarity  38.5%;  Pred. No. 1.6e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 TGEKAMMLLGQVK 15
          | |    :|||::
Db      3 TSEGCRQILGQLQ 15


RESULT  14
5304631-8
;Patent No. 5304631
;    APPLICANT: STEWART, JOHN M.;HAHN, KARL W.;. KLIS, WIESLAW A.
;    TITLE OF INVENTION: SYNTHETIC HELIZYME ENZYMES
;    NUMBER OF SEQUENCES: 16
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/464,932
;      FILING DATE: 16-JAN-1990
;SEQ ID NO:8:
;      LENGTH: 15
5304631-8


  Query Match           25.0%;  Score 26;  DB 5;  Length 15;
  Best Local Similarity  54.5%;  Pred. No. 1.6e+02;
```

```
;       OPERATING SYSTEM:  Macintosh 6.0.7
;       SOFTWARE:  Microsoft Word 4.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/920,430
;       FILING DATE:  19920731
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/796,256
;       FILING DATE:  20 NOV 1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/767,251
;       FILING DATE:  27 SEPT 1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/659,975
;       FILING DATE:  22 FEB 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Elizabeth Lassen
;       REGISTRATION NUMBER:  31,845
;       NAME:  Donna E. Scherer
;       REGISTRATION NUMBER:  34,719
;       REFERENCE/DOCKET NUMBER:  CGNE 79-3
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (916) 753-6313
;       TELEFAX:  (916) 753-1510
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-07-920-430-3


  Query Match             25.0%;  Score 26;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DITGE 5
           |||||
Db       7 DITGE 11


RESULT  10
US-08-251-464-3
; Sequence 3, Application US/08251464
; Patent No. 5723747
;  GENERAL INFORMATION:
;     APPLICANT:  Lassner, Michael
;     APPLICANT:  Metz, James
;     TITLE OF INVENTION:  WAX ESTERS IN TRANSFORMEDPLANTS
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Calgene, Inc.
;       STREET:  1920 Fifth Street
;       CITY:  Davis
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  95616
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 MB
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  Macintosh 7.0
;       SOFTWARE:  Microsoft Word 5.1a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/251,464
;       FILING DATE:  01-JUN-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Elizabeth Lassen
;       REGISTRATION NUMBER:  31,845
;       NAME:  Donna E. Scherer
;       REGISTRATION NUMBER:  34,719
;       NAME:  Carl J. Schwedler
;       REGISTRATION NUMBER:  36,924
;       REFERENCE/DOCKET NUMBER:  CGNE 107
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (916) 753-6313
;       TELEFAX:  (916) 753-1510
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-251-464-3


  Query Match             25.0%;  Score 26;  DB 1;  Length 13;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DITGE 5
           |||||
Db       7 DITGE 11


RESULT  11
PCT-US92-01364-3
; Sequence 3, Application PC/TUS9201364
;  GENERAL INFORMATION:
;     APPLICANT:  Calgene, Inc.
;     TITLE OF INVENTION:  Fatty Acyl Reductase
;     NUMBER OF SEQUENCES:  19
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Calgene, Inc.
;       STREET:  1920 Fifth Street
;       CITY:  Davis
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  95616
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 MB
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  Macintosh 6.0.7
;       SOFTWARE:  Microsoft Word 4.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/01364
;       FILING DATE:  19920212
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/659,975
;       FILING DATE:  22 FEB 1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/767,251
;       FILING DATE:  27 SEPT 1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/796,256
;       FILING DATE:  20 NOV 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Elizabeth Lassen
;       REGISTRATION NUMBER:  31,845
;       NAME:  Donna E. Scherer
;       REGISTRATION NUMBER:  34,719
;       REFERENCE/DOCKET NUMBER:  CGNE79-2WO
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (916) 753-6313
;       TELEFAX:  (916) 753-1510
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  13 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
PCT-US92-01364-3
```

```
;    MOLECULE TYPE:  peptide
US-08-355-888A-25


  Query Match             26.0%;  Score 27;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 91;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   7
US-08-693-697-25
; Sequence 25, Application US/08693697
; Patent No. 5869610
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,697
;      FILING DATE:  05-AUG-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0037-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-693-697-25


  Query Match             26.0%;  Score 27;  DB 2;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 91;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   8
US-08-693-696-25
; Sequence 25, Application US/08693696
; Patent No. 6005080
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,696
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/355,888
;      FILING DATE:  14-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-078
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-693-696-25


  Query Match             26.0%;  Score 27;  DB 3;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 91;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   9
US-07-920-430-3
; Sequence 3, Application US/07920430
; Patent No. 5370996
;  GENERAL INFORMATION:
;    APPLICANT:  James George Metz
;    APPLICANT:  Michael Roman Pollard
;    APPLICANT:  Michael W. Lassner
;    TITLE OF INVENTION:  Fatty Acyl Reductases
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Calgene, Inc.
;      STREET:  1920 Fifth Street
;      CITY:  Davis
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  95616
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.0 MB
;      COMPUTER:  Apple Macintosh
```

```
;       FILING DATE:  06-SEP-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/689,807
;       FILING DATE:  14-AUG-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/537,400
;       FILING DATE:  02-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/370,190
;       FILING DATE:  01-JAN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/316,423
;       FILING DATE:  30-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/576,759
;       FILING DATE:  21-DEC-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/537,397
;       FILING DATE:  02-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/525,057
;       FILING DATE:  08-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Esmond, Robert W.
;       REGISTRATION NUMBER:  32,893
;       REFERENCE/DOCKET NUMBER:  0942.2800005
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-2600
;       TELEFAX:  202-371-2540
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  not relevant
;     MOLECULE TYPE:  peptide
US-08-706-706-21


  Query Match             26.0%;  Score 27;  DB 3;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 68;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMM 9
            |||:| |:
Db       3 ITGDKDML 10


RESULT   5
US-08-306-231-9
; Sequence 9, Application US/08306231
; Patent No. 5643748
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,231
;      FILING DATE:  14-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-076
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-306-231-9


  Query Match             26.0%;  Score 27;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 91;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   6
US-08-355-888A-25
; Sequence 25, Application US/08355888A
; Patent No. 5763211
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/355,888A
;      FILING DATE:  14-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-078
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
```

```
Qy      14 VKYGLHN 20
           ||:| ||
Db       3 VKHGSHN 9


RESULT   2
US-08-159-340A-17
; Sequence 17, Application US/08159340A
; Patent No. 5565352
;  GENERAL INFORMATION:
;    APPLICANT:  Hochstrasser, Mark
;    APPLICANT:  Papa, Feroz
;    TITLE OF INVENTION:  DEUBIQUITINATING ENZYME:  COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,340A
;      FILING DATE:  24-NOV-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  ARCD:112/HYL
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-340A-17


  Query Match             27.9%;  Score 29;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 VKYGLHN 20
           | || ||
Db       8 VHYGTHN 14


RESULT   3
US-08-706-702-21
; Sequence 21, Application US/08706702
; Patent No. 5948614
;  GENERAL INFORMATION:
;    APPLICANT:  Chatterjee, Deb K.
;    TITLE OF INVENTION:  Cloned DNA Polymerases from Thermotoga
;    TITLE OF INVENTION:  maritima and Mutants Thereof
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 New York Ave., N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/706,702
;      FILING DATE:  06-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/689,807
;      FILING DATE:  14-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/537,400
;      FILING DATE:  02-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/576,759
;      FILING DATE:  21-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/537,397
;      FILING DATE:  02-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/525,057
;      FILING DATE:  08-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Esmond, Robert W.
;      REGISTRATION NUMBER:  32,893
;      REFERENCE/DOCKET NUMBER:  0942.2800006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-706-702-21


  Query Match             26.0%;  Score 27;  DB 2;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 68;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMM 9
            |||:| |:
Db       3 ITGDKDML 10


RESULT   4
US-08-706-706-21
; Sequence 21, Application US/08706706
; Patent No. 6015668
;  GENERAL INFORMATION:
;    APPLICANT:  Hughes, A. John
;    APPLICANT:  Chatterjee, Deb K.
;    TITLE OF INVENTION:  Cloned DNA Polymerases from Thermotoga and
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 New York Ave., N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/706,706
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:49 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-13
Perfect score: 104
Sequence: 1 DITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	27.9	12	1	US-08-164-618-20	Sequence 20, Appl
2	29	27.9	14	1	US-08-159-340A-17	Sequence 17, Appl
3	27	26.0	10	2	US-08-706-702-21	Sequence 21, Appl
4	27	26.0	10	3	US-08-706-706-21	Sequence 21, Appl
5	27	26.0	13	1	US-08-306-231-9	Sequence 9, Appli
6	27	26.0	13	1	US-08-355-888A-25	Sequence 25, Appl
7	27	26.0	13	2	US-08-693-697-25	Sequence 25, Appl
8	27	26.0	13	3	US-08-693-696-25	Sequence 25, Appl
9	26	25.0	13	1	US-07-920-430-3	Sequence 3, Appli
10	26	25.0	13	1	US-08-251-464-3	Sequence 3, Appli
11	26	25.0	13	4	PCT-US92-01364-3	Sequence 3, Appli
12	26	25.0	15	2	US-08-484-905-28	Sequence 28, Appl
13	26	25.0	15	3	US-08-481-985B-28	Sequence 28, Appl
14	26	25.0	15	5	5304631-8	Patent No. 5304631
15	26	25.0	15	5	5304631-16	Patent No. 5304631
16	26	25.0	17	2	US-08-182-067-14	Sequence 14, Appl
17	26	25.0	17	2	US-08-465-313-14	Sequence 14, Appl
18	26	25.0	17	5	5304631-5	Patent No. 5304631
19	26	25.0	20	2	US-08-466-975A-12	Sequence 12, Appl
20	26	25.0	20	2	US-08-466-975A-13	Sequence 13, Appl
21	26	25.0	20	2	US-08-391-671A-12	Sequence 12, Appl
22	26	25.0	20	2	US-08-391-671A-13	Sequence 13, Appl
23	26	25.0	20	2	US-08-475-634D-2	Sequence 2, Appli
24	26	25.0	20	3	US-08-467-902A-12	Sequence 12, Appl
25	26	25.0	20	3	US-08-467-902A-13	Sequence 13, Appl
26	25	24.0	15	2	US-08-480-190-77	Sequence 77, Appl
27	25	24.0	15	2	US-08-488-379-77	Sequence 77, Appl
28	25	24.0	15	4	PCT-US93-07545-77	Sequence 77, Appl
29	25	24.0	16	2	US-08-480-190-76	Sequence 76, Appl
30	25	24.0	16	2	US-08-488-379-76	Sequence 76, Appl
31	25	24.0	16	4	PCT-US93-07545-76	Sequence 76, Appl
32	25	24.0	17	1	US-07-894-212A-3	Sequence 3, Appli
33	25	24.0	18	1	US-07-893-928A-3	Sequence 3, Appli
34	25	24.0	18	2	US-08-480-190-192	Sequence 192, App
35	25	24.0	18	2	US-08-488-379-192	Sequence 192, App
36	25	24.0	18	4	PCT-US93-07545-192	Sequence 192, App
37	25	24.0	19	2	US-08-480-190-74	Sequence 74, Appl
38	25	24.0	19	2	US-08-480-190-75	Sequence 75, Appl
39	25	24.0	19	2	US-08-488-379-74	Sequence 74, Appl
40	25	24.0	19	2	US-08-488-379-75	Sequence 75, Appl
41	25	24.0	19	2	US-08-696-944-3	Sequence 3, Appli
42	25	24.0	19	4	PCT-US93-07545-74	Sequence 74, Appl
43	25	24.0	19	4	PCT-US93-07545-75	Sequence 75, Appl
44	25	24.0	20	2	US-08-480-190-191	Sequence 191, App
45	25	24.0	20	2	US-08-749-852-55	Sequence 55, Appl

ALIGNMENTS

```
RESULT 1
US-08-164-618-20
; Sequence 20, Application US/08164618
; Patent No. 5408036
;  GENERAL INFORMATION:
;    APPLICANT:  Ghadiri, M. Reza
;    TITLE OF INVENTION:  Isolated Metallopolypeptides:
;    TITLE OF INVENTION:  Compositions and Synthetic Methods
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dressler, Goldsmith, Shore, Sutker &
;      ADDRESSEE:  Milnamow, Ltd.
;      STREET:  180 No. 5408036th Stetson, Suite 4700
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/164,618
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/769,621
;      FILING DATE:
;      APPLICATION NUMBER:  07/591,988
;      FILING DATE:  October 2, 1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gamson, Edward G.
;      REGISTRATION NUMBER:  29,381
;      REFERENCE/DOCKET NUMBER:  SCRF 231.0
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 616-5400
;      TELEFAX:  (312) 616-5460
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-164-618-20


  Query Match             27.9%;  Score 29;  DB 1;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 37;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
RT   Helicobacter species from humans and animals.";
RL   Infect. Immun. 60:5259-5266(1992).
SQ   SEQUENCE   18 AA;  2060 MW;  29C8E0AB77E21805 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGE 8
          ::||  ||
Db      9 MMLHYAGE 16


RESULT  12
Q9PRM7
ID   Q9PRM7      PRELIMINARY;      PRT;    18 AA.
AC   Q9PRM7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ACID THIOL PROTEASE (FRAGMENT).
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96417001.
RA   Miyata S., Kihara H.K.;
RT   "Cathepsin L-like protease from Xenopus embryos that is stimulated by
RT   nucleoside phosphates and nucleic acids.";
RL   Zool. Sci. 12:771-774(1995).
SQ   SEQUENCE   18 AA;  2060 MW;  A1DC9B106B341504 CRC64;


  Query Match             22.9%;  Score 25;  DB 13;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 REPGWIK 15
          || |::|
Db      8 REKGYVK 14


RESULT  13
Q44850
ID   Q44850      PRELIMINARY;      PRT;    19 AA.
AC   Q44850;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   HYPOTHETICAL PROTEIN (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RA   Akins D.R., Popova T., Brusca J., Goldberg M.L., Li M., Baker S.C.,
RA   Norgard M.V., Radolf J.D.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L31425; AAA64901.1; -.
KW   Hypothetical protein.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2321 MW;  C821BF312DB41512 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 LLHLQGEREP 11
          |:|||    |
Db      4 LIHLQSILHP 13


RESULT  14
Q9UJI8
ID   Q9UJI8      PRELIMINARY;      PRT;    19 AA.
AC   Q9UJI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ333A15.2 (PROSTAGLANDIN E RECEPTOR (EP3E)) (FRAGMENT).
GN   PTGER3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Frankland J.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031429; CAB52457.1; -.
KW   Receptor.
FT   NON_TER      1      1
SQ   SEQUENCE   19 AA;  2245 MW;  142450379B298FEC CRC64;


  Query Match             22.9%;  Score 25;  DB 4;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 LHLQGEREP 11
          ||: |: :|
Db      8 LHIVGKYKP 16


RESULT  15
O12088
ID   O12088      PRELIMINARY;      PRT;    12 AA.
AC   O12088;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.F., Vigne R., Querat G.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81435; AAB60824.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   12 AA;  1279 MW;  4B90BBB1E8644EB7 CRC64;


  Query Match             22.0%;  Score 24;  DB 12;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     11 PGW 13
          |||
Db      8 PGW 10


Search completed: December 21, 2000, 08:35:39
Job time: 602 sec
```

```
Qy      13 WIKQLF 18
           ||::|:
Db       1 WIRELY 6


RESULT   7
Q9X3I3
ID   Q9X3I3      PRELIMINARY;      PRT;    17 AA.
AC   Q9X3I3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B (FRAGMENT).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070176; AAD20791.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   17 AA;  1953 MW;  630843039ADD51B4 CRC64;

  Query Match             23.9%;  Score 26;  DB 2;  Length 17;
  Best Local Similarity   33.3%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGEREPG 12
           :|:|   :|: |
Db       1 MLMHFLXDRKQG 12


RESULT   8
Q9T2R9
ID   Q9T2R9      PRELIMINARY;      PRT;    17 AA.
AC   Q9T2R9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOCHROME-C REDUCTASE 55 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS   Solanum tuberosum (Potato).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RA   Braun H.P., Kruft V., Schmitz U.K.;
RL   Planta 193:99-106(1994).
SQ   SEQUENCE   17 AA;  1758 MW;  F95F7BF7940F5F21 CRC64;

  Query Match             23.9%;  Score 26;  DB 8;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LLLHLQG 7
           |:||: |
Db       4 LMLHIDG 10


RESULT   9
Q9UCG3
ID   Q9UCG3      PRELIMINARY;      PRT;    19 AA.
AC   Q9UCG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ALPHA 2-PLASMIN INHIBITOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93249387.
RA   Bangert K., Johnsen A.H., Christensen U., Thorsen S.;
RT   "Different N-terminal forms of alpha 2-plasmin inhibitor in human
RT   plasma.";
RL   Biochem. J. 291:623-625(1993).
SQ   SEQUENCE   19 AA;  2065 MW;  41352BF04D1EEAE9 CRC64;

  Query Match             23.4%;  Score 25.5;  DB 4;  Length 19;
  Best Local Similarity   63.6%;  Pred. No. 1.5e+03;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       2 LLHLQGEREPG 12
           || | | :|||
Db      10 LLKL-GNQEPG 19


RESULT  10
Q16045
ID   Q16045      PRELIMINARY;      PRT;    14 AA.
AC   Q16045;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   D3 DOPAMINE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93326145.
RA   Nagai Y., Ueno S., Saeki Y., Soga F., Yanagihara T.;
RT   "Expression of the D3 dopamine receptor gene and a novel variant
RT   transcript generated by alternative splicing in human peripheral blood
RT   lymphocytes.";
RL   Biochem. Biophys. Res. Commun. 194:368-374(1993).
DR   EMBL; S63845; CAB32270.1; -.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1586 MW;  EA310BEFE94CF1B1 CRC64;

  Query Match             22.9%;  Score 25;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 LLHLQ 6
           |||||
Db       9 LLHLQ 13


RESULT  11
Q9R5F6
ID   Q9R5F6      PRELIMINARY;      PRT;    18 AA.
AC   Q9R5F6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   UREASE SMALL SUBUNIT (FRAGMENT).
OS   Helicobacter mustelae.
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93084378.
RA   Turbett G.R., Hoj P.B., Horne R., Mee B.J.;
RT   "Purification and characterization of the urease enzymes of
```

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95337087.
RA   Thulin C.D., Walsh K.A.;
RT   "Identification of the amino terminus of human filaggrin using
RT   differential LC/MS techniques: implications for profilaggrin
RT   processing.";
RL   Biochemistry 34:8687-8692(1995).
SQ   SEQUENCE   20 AA;  2199 MW;  0FD0856B199AB332 CRC64;


  Query Match             26.6%;  Score 29;  DB 4;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 4.6e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      4 HLQGEREPGW 13
          | | |   ||
Db      5 HEQSESSHGW 14


RESULT   3
Q9PRJ4
ID   Q9PRJ4      PRELIMINARY;      PRT;     9 AA.
AC   Q9PRJ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BRADYKININ.
OS   Lepisosteus osseus (Long-nosed gar), and Amia calva (Bowfin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC   Lepisosteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95380361.
RA   Conlon J.M., Platzack B., Marra L.E., Youson J.H., Olson K.R.;
RT   "Isolation and biological activity of [Trp5]bradykinin from the plasma
RT   of the phylogenetically ancient fish, the bowfin and the longnosed
RT   gar.";
RL   Peptides 16:485-489(1995).
SQ   SEQUENCE   9 AA;  1099 MW;  3393D775A3786777 CRC64;


  Query Match             25.7%;  Score 28;  DB 13;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 REPGW 13
          | |||
Db      1 RPPGW 5


RESULT   4
O18502
ID   O18502      PRELIMINARY;      PRT;    14 AA.
AC   O18502;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   LYSOPHOSPHOLIPASE HOMOLOG (FRAGMENT).
GN   SMLPLH.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGYPTIAN;
RA   Hamdan F.F., Ribeiro P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF006679; AAC62255.1; -.
FT   NON_TER      14      14
SQ   SEQUENCE   14 AA;  1541 MW;  48B847C2E5D89177 CRC64;


  Query Match             25.7%;  Score 28;  DB 5;  Length 14;
  Best Local Similarity   38.5%;  Pred. No. 4.5e+02;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGEREPGW 13
            : ||  |:   ||
Db      2 IFLHGLGDTGHGW 14


RESULT   5
Q9R4W5
ID   Q9R4W5      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4W5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   30 KDA MAJOR HEAT SHOCK PROTEIN (FRAGMENT).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95020803.
RA   Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,
RA   Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
RT   "Heat shock protein produced by Helicobacter pylori.";
RL   Microbiol. Immunol. 38:403-405(1994).
DR   INTERPRO; IPR002026; -.
DR   PFAM; PF00547; urease_gamma; 1.
SQ   SEQUENCE   20 AA;  2302 MW;  29C9DFBFD6D21805 CRC64;


  Query Match             24.8%;  Score 27;  DB 2;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 9.4e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGE 8
            |:||  ||
Db     11 LMLHYAGE 18


RESULT   6
O77893
ID   O77893      PRELIMINARY;      PRT;    11 AA.
AC   O77893;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MHC CLASS II B LOCUS 10 (FRAGMENT).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC   Percomorpha; Perciformes; Labroidei; Cichlidae; Oreochromis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98315113.
RA   Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA   Sueltmann H., Figueroa F., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT   class II B loci.";
RL   Genetics 149:1527-1547(1998).
DR   EMBL; AF050003; AAC41342.1; -.
KW   MHC.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1296 MW;  68775B73786B572B CRC64;


  Query Match             23.9%;  Score 26;  DB 7;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 7.3e+02;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:39 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-12
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	26.6	10	13	Q9PRZ1	Q9prz1 oncorhynchu
2	29	26.6	20	4	Q9UC71	Q9uc71 homo sapien
3	28	25.7	9	13	Q9PRJ4	Q9prj4 lepisosteus
4	28	25.7	14	5	O18502	O18502 schistosoma
5	27	24.8	20	2	Q9R4W5	Q9r4w5 helicobacte
6	26	23.9	11	7	O77893	O77893 oreochromis
7	26	23.9	17	2	Q9X3I3	Q9x3i3 prochloroco
8	26	23.9	17	8	Q9T2R9	Q9t2r9 solanum tub
9	25.5	23.4	19	4	Q9UCG3	Q9ucg3 homo sapien
10	25	22.9	14	4	Q16045	Q16045 homo sapien
11	25	22.9	18	2	Q9R5F6	Q9r5f6 helicobacte
12	25	22.9	18	13	Q9PRM7	Q9prm7 xenopus. ac
13	25	22.9	19	2	Q44850	Q44850 borrelia bu
14	25	22.9	19	4	Q9UJI8	Q9uji8 homo sapien
15	24	22.0	12	12	O12088	O12088 caprine art
16	24	22.0	12	12	O12074	O12074 caprine art
17	24	22.0	12	12	O12076	O12076 caprine art
18	24	22.0	12	12	O12078	O12078 caprine art
19	24	22.0	12	12	O12080	O12080 caprine art
20	24	22.0	12	12	O12082	O12082 caprine art
21	24	22.0	12	12	O12084	O12084 caprine art
22	24	22.0	12	12	O12086	O12086 caprine art
23	24	22.0	12	12	O12090	O12090 caprine art
24	24	22.0	12	12	O12092	O12092 caprine art
25	24	22.0	12	12	O12094	O12094 caprine art
26	24	22.0	12	12	O12106	O12106 caprine art
27	24	22.0	12	12	O12108	O12108 caprine art
28	24	22.0	12	12	O12110	O12110 caprine art
29	24	22.0	12	12	O12112	O12112 caprine art
30	24	22.0	12	12	O12114	O12114 caprine art
31	24	22.0	12	12	O12116	O12116 caprine art
32	24	22.0	12	12	O12118	O12118 caprine art
33	24	22.0	15	2	Q46963	Q46963 escherichia
34	24	22.0	15	5	Q9TXC8	Q9txc8 locusta mig
35	24	22.0	16	4	Q9UCK9	Q9uck9 homo sapien
36	24	22.0	16	4	Q9UCH1	Q9uch1 homo sapien
37	24	22.0	19	2	Q47079	Q47079 escherichia
38	24	22.0	19	10	Q9S8W5	Q9s8w5 avena sativ
39	24	22.0	19	12	O93210	O93210 porcine cir
40	23	21.1	11	5	Q9V7K6	Q9v7k6 drosophila
41	23	21.1	11	7	O77892	O77892 oreochromis
42	23	21.1	11	7	O77906	O77906 oreochromis
43	23	21.1	11	7	O77918	O77918 pseudotroph
44	23	21.1	13	4	Q9Y674	Q9y674 homo sapien
45	23	21.1	17	6	Q9TR78	Q9tr78 didelphis m

ALIGNMENTS

RESULT 1
Q9PRZ1
ID Q9PRZ1 PRELIMINARY; PRT; 10 AA.
AC Q9PRZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE LYSYL-BRADYKININ=VASOACTIVE PEPTIDE KALLIDIN HOMOLOG.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 94039817.
RA Conlon J.M., Olson K.R.;
RT "Purification of a vasoactive peptide related to lysyl-bradykinin from trout plasma.";
RL FEBS Lett. 334:75-78(1993).
SQ SEQUENCE 10 AA; 1193 MW; 33C59075A3786777 CRC64;

Query Match 26.6%; Score 29; DB 13; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 EREPGW 13
:| |||
Db 1 KRPPGW 6

RESULT 2
Q9UC71
ID Q9UC71 PRELIMINARY; PRT; 20 AA.
AC Q9UC71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE FILAGGRIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
CC   -!- FUNCTION: SHOWS HEMOLYTIC ACTIVITY.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: SKIN.
DR   PIR; S09018; S09018.
KW   Amphibian skin; Amidation; Hemolysis.
FT   MOD_RES       13      13       AMIDATION.
FT   NON_TER       13      13
SQ   SEQUENCE   13 AA;  1390 MW;  C6BA768B9DFE587D CRC64;


  Query Match             18.3%;  Score 20;  DB 1;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
             | : |||
Db       7 GILSQLF 13


Search completed: December 21, 2000, 08:32:59
Job time: 442 sec
```

```
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85083111.
RA   Rodakis G.C., Lecanidou R., Eickbush T.H.;
RT   "Diversity in a chorion multigene family created by tandem
RT   duplications and a putative gene-conversion event.";
RL   J. Mol. Evol. 20:265-273(1984).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC       SILK MOTH.
CC   -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
CC       BELONG CLASSES A, CA AND HCA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01068; CAB57790.1; -.
DR   PIR; B23219; B23219.
KW   Eggshell; Chorion; Repeat; Multigene family; Signal.
FT   SIGNAL        1    >17
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1913 MW;  5E634508C5355C9C CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 17;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LLLHLQG 7
           ||| :||
Db       7 LLLCVQG 13


RESULT  13
CRTC_SPIOL
ID   CRTC_SPIOL     STANDARD;      PRT;    20 AA.
AC   P30806;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   CALRETICULIN-LIKE PROTEIN (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LEAF;
RX   MEDLINE; 93176159.
RA   Menegazzi P., Guzzo F., Baldan B., Mariani P., Treves S.;
RT   "Purification of calreticulin-like protein(s) from spinach leaves.";
RL   Biochem. Biophys. Res. Commun. 190:1130-1135(1993).
CC   -!- FUNCTION: THIS PROTEIN BINDS CALCIUM.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
DR   PIR; PC1241; PC1241.
DR   PIR; PC1240; PC1240.
DR   INTERPRO; IPR001580; -.
DR   PROSITE; PS00803; CALRETICULIN_1; PARTIAL.
DR   PROSITE; PS00804; CALRETICULIN_2; PARTIAL.
DR   PROSITE; PS00805; CALRETICULIN_REPEAT; PARTIAL.
KW   Endoplasmic reticulum; Calcium-binding; Glycoprotein.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2645 MW;  00FAB4C9DEEDCB0F CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   75.0%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 EPGW 13
             | ||
Db       9 EDGW 12


RESULT  14
PORD_METTM
ID   PORD_METTM     STANDARD;      PRT;    12 AA.
AC   P80903;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PYRUVATE SYNTHASE SUBUNIT PORD (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE
DE   DELTA CHAIN) (POR) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE DELTA SUBUNIT)
DE   (FRAGMENT).
GN   PORD.
OS   Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC   Methanobacterium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 97261844.
RA   Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT   "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT   in Methanobacterium thermoautotrophicum.";
RL   Eur. J. Biochem. 244:862-868(1997).
CC   -!- FUNCTION: THE PH OPTIMUM IS PH 10.0 AND THE OPTIMAL TEMPERATURE
CC       IS 80 DEGREES CELSIUS.
CC   -!- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN =
CC       ACETYL-COA + CO(2) + REDUCED FERREDOXIN.
CC   -!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (BY SIMILARITY).
CC   -!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE
CC       GAMMA CHAIN.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
KW   Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1241 MW;  2D54065D1BD1ADD8 CRC64;


  Query Match             18.3%;  Score 20;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 REPG 12
           :|||
Db       9 KEPG 12


RESULT  15
HPA1_RANES
ID   HPA1_RANES     STANDARD;      PRT;    13 AA.
AC   P32415;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   HEMOLYTIC PROTEIN A1 (FRAGMENT).
OS   Rana esculenta (Edible frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN SECRETION;
RX   MEDLINE; 90198965.
RA   Simmaco M., de Biase D., Severini C., Aita M., Erspamer G.F.,
RA   Barra D., Bossa F.;
RT   "Purification and characterization of bioactive peptides from skin
RT   extracts of Rana esculenta.";
RL   Biochim. Biophys. Acta 1033:318-323(1990).
```

```
CC   -!- FUNCTION: SMOOTH MUSCLE CONTRACTION.
KW   Bradykinin; Vasodilator.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1573 MW;  2673CB3D83ECC867 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 QGEREPGW 13
          | :| ||:
Db      1 QZKRPPGF 8


RESULT   9
NEUT_CHICK
ID   NEUT_CHICK     STANDARD;      PRT;    13 AA.
AC   P13724;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   NEUROTENSIN (NT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88063566.
RA   Iwabuchi H., Komori S., Ohashi H., Kimura S.;
RT   "The amino acid sequence of a smooth muscle-contracting peptide from
RT   chicken rectum. Identity to chicken neurotensin.";
RL   Jpn. J. Pharmacol. 44:455-459(1987).
CC   -!- FUNCTION: SMOOTH MUSCLE-CONTRACTING PEPTIDE.
DR   PIR; A28505; A28505.
KW   Vasoactive.
FT   MOD_RES       1       1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1608 MW;  4C949E714C410DD3 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 LHLQGEREP 11
          ||:   | |
Db      2 LHVNKARRP 10


RESULT  10
Q2OG_COMTE
ID   Q2OG_COMTE     STANDARD;      PRT;    10 AA.
AC   P80466;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   QUINOLINE 2-OXIDOREDUCTASE, GAMMA CHAIN (EC 1.-.-.-) (FRAGMENT).
OS   Comamonas testosteroni (Pseudomonas testosteroni).
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=63;
RX   MEDLINE; 96035889.
RA   Schach S., Tshisuaka B., Fetzner S., Lingens F.;
RT   "Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT   dioxygenase from Comamonas testosteroni 63. The first two enzymes in
RT   quinoline and 3-methylquinoline degradation.";
RL   Eur. J. Biochem. 232:536-544(1995).
CC   -!- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)2-OXO-
CC       1,2-DIHYDROQUINOLINE.
CC   -!- COFACTOR: FAD AND MOLYBDENUM.
CC   -!- PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
CC       (3-METHYL-)QUINOLINE.
CC   -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC       TWO GAMMA CHAINS (PROBABLE).
KW   Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT   NON_TER      10      10
SQ   SEQUENCE   10 AA;  1153 MW;  C848CE64433B1DC6 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   42.9%;  Pred. No. 1e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 LQGEREP 11
          :| |: |
Db      2 IQAEKNP 8


RESULT  11
CXET_CONTE
ID   CXET_CONTE     STANDARD;      PRT;    13 AA.
AC   P81755;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   EPSILON-CONOTOXIN TXIX.
OS   Conus textile (Cloth-of-gold cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
RN   [1]
RP   SEQUENCE, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY NMR.
RC   TISSUE=VENOM;
RX   MEDLINE; 99254114.
RA   Rigby A.C., Lucas-Meunier E., Kalume D.E., Czerwiec E., Hambe B.,
RA   Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,
RA   Furie B.C., Furie B., Stenflo J.P.;
RT   "A conotoxin from Conus textile with unusual posttranslational
RT   modifications reduces presynaptic Ca2+ influx.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).
CC   -!- FUNCTION: CONOTOXIN WHICH ACTS AT PRESYNAPTIC MEMBRANES, BLOCKING
CC       THE CALCIUM CHANNELS.
CC   -!- PTM: O-GLYCAN CONSISTS OF THE DISACCHARIDE GAL-GALNAC.
DR   PDB; 1WCT; 08-JUN-99.
KW   Presynaptic neurotoxin; Calcium channel inhibitor; Venom; Vitamin K;
KW   Gamma-carboxyglutamic acid; Glycoprotein; Hydroxylation; Bromination;
KW   3D-structure.
FT   DISULFID      2       8
FT   DISULFID      3       9
FT   MOD_RES       1       1       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       4       4       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       7       7       BROMINATION.
FT   MOD_RES      13      13       HYDROXYLATION.
FT   CARBOHYD     10      10       O-LINKED (GALNAC. . .).
SQ   SEQUENCE   13 AA;  1388 MW;  386C9E1C74AFA378 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     10 EPGW 13
          | ||
Db      4 EDGW 7


RESULT  12
CHH3_BOMMO
ID   CHH3_BOMMO     STANDARD;      PRT;    17 AA.
AC   P20729;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   CHORION CLASS HIGH-CYSTEINE HCA PROTEIN 13 PRECURSOR (HC-A.13)
DE   (FRAGMENT).
```

```
DE   (DCC II).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 90046758.
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; B33995; B33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1169 MW;  916036786771A9D1 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   6
AROQ_AMYME
ID   AROQ_AMYME     STANDARD;      PRT;    20 AA.
AC   P46380;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10) (3-DEHYDROQUINASE)
DE   (TYPE II DHQASE) (FRAGMENT).
GN   AROQ.
OS   Amycolatopsis methanolica.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC   Amycolatopsis.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=NCIB 11946;
RX   MEDLINE; 93123995.
RA   Euverink G.J.W., Hessels G.I., Vrijbloed J.W., Coggins J.R.,
RA   Dijkhuizen L.;
RT   "Purification and characterization of a dual function
RT   3-dehydroquinate dehydratase from Amycolatopsis methanolica.";
RL   J. Gen. Microbiol. 138:2449-2457(1992).
CC   -!- FUNCTION: CATALYZE A TRANS-DEHYDRATION VIA AN ENOLATE
CC       INTERMEDIATE. IS INVOLVED IN BOTH THE CATABOLISM OF QUINATE AND
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS. HAS A TEMPERATURE
CC       OPTIMAL OF 76 DEGREES CELSIUS AND PH OPTIMAL OF 9.0. TYPE II
CC       ENZYMES ARE THERMOSTABLE.
CC   -!- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE = 3-DEHYDROSHIKIMATE + H(2)O.
CC   -!- PATHWAY: THIRD STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: HOMODODECAMER.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-II 3-DEHYDROQUINASE FAMILY.
DR   INTERPRO; IPR001874; -.
DR   PROSITE; PS01029; DEHYDROQUINASE_II; PARTIAL.
KW   Quinate metabolism; Aromatic amino acid biosynthesis; Lyase.
FT   NON_TER       20     20
SQ   SEQUENCE   20 AA;  2197 MW;  C24AA183E5CFDF0A CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GEREP 11
            |:|||
Db      15 GKREP 19


RESULT   7
LPK1_LOCMI
ID   LPK1_LOCMI     STANDARD;      PRT;    16 AA.
AC   P20404;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   LOCUSTAPYROKININ 1 (LOM-PK-1).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 91224474.
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, primary structure, and synthesis of locustapyrokinin: a
RT   myotropic peptide of Locusta migratoria.";
RL   Gen. Comp. Endocrinol. 81:97-104(1991).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   PIR; A49761; A49761.
DR   INTERPRO; IPR001484; -.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1827 MW;  A7178BBDCA0AFDD6 CRC64;


  Query Match             21.1%;  Score 23;  DB 1;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 8e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           || :| |
Db       6 GWPQQPF 12


RESULT   8
BRK_PARID
ID   BRK_PARID      STANDARD;      PRT;    13 AA.
AC   P42717;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   WASPKININ.
OS   Parapolybia indica.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Polistinae; Parapolybia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Toki T., Yasuhara T., Nakajima T.;
RT   "Isolation and sequential analysis of peptides on the venom sac of
RT   Parapolybia indica.";
RL   Eisei Dobutsu 39:105-111(1988).
```

```
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  86786771A9D1A736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   3
RPCH_PANBO
ID   RPCH_PANBO     STANDARD;      PRT;     8 AA.
AC   P08939;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RED PIGMENT CONCENTRATING HORMONE (RPCH).
OS   Pandalus borealis (Northern red shrimp).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidea;
OC   Pandalidae; Pandalus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 75054965.
RA   Fernlund P.;
RT   "Structure of the red-pigment-concentrating hormone of the shrimp,
RT   Pandalus borealis.";
RL   Biochim. Biophys. Acta 371:304-311(1974).
CC   -!- FUNCTION: THIS HORMONE ADAPTS THE ANIMAL TO LIGHT BACKGROUNDS BY
CC       STIMULATING CONCENTRATION OF THE PIGMENT OF ITS RED BODY-
CC       CHROMATOPHORES.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S07139; S07139.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Pigment; Hormone; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  948 MW;  86786775B9C44736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   4
HTF_NAUCI
ID   HTF_NAUCI      STANDARD;      PRT;    10 AA.
AC   P10939;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPERTREHALOSAEMIC HORMONE (HTH) (HYPERTREHALOSAEMIC NEUROPEPTIDE).
OS   Nauphoeta cinerea (Cinereous cockroach) (Gray cockroach),
OS   Leucophaea maderae (Madeira cockroach),
OS   Blattella germanica (German cockroach), and
OS   Gromphadorina portentosa (Cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Nauphoeta.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=N.CINEREA; TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 87100208.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Amino acid sequence of a hypertrehalosaemic neuropeptide from the
RT   corpus cardiacum of the cockroach, Nauphoeta cinerea.";
RL   Biochem. Biophys. Res. Commun. 141:774-781(1986).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=L.MADERAE, G.PORTENTOSA, AND B.GERMANICA;
RX   MEDLINE; 90253659.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=B.GERMANICA;
RX   MEDLINE; 91179584.
RA   Veenstra J.A., Camps F.;
RT   "Structure of the hypertrehalosemic neuropeptide of the German
RT   cockroach, Blattella germanica.";
RL   Neuropeptides 15:107-109(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A26381; A26381.
DR   PIR; S08997; S08997.
DR   PIR; S08998; S08998.
DR   PIR; S09137; S09137.
DR   PIR; A60421; A60421.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1092 MW;  056236786775B9C4 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   5
HTF_TABAT
ID   HTF_TABAT      STANDARD;      PRT;    10 AA.
AC   P14596;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   HYPERTREHALOSAEMIC FACTOR (HOTH) (DIPTERAN CORPORA CARDIACA FACTOR II)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:59 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-12
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.7	19	1	TRP3_LEUMA	P81735 leucophaea
2	24	22.0	8	1	AKH_TABAT	P14595 tabanus atr
3	24	22.0	8	1	RPCH_PANBO	P08939 pandalus bo
4	24	22.0	10	1	HTF_NAUCI	P10939 nauphoeta c
5	24	22.0	10	1	HTF_TABAT	P14596 tabanus atr
6	24	22.0	20	1	AROQ_AMYME	P46380 amycolatops
7	23	21.1	16	1	LPK1_LOCMI	P20404 locusta mig
8	22	20.2	13	1	BRK_PARID	P42717 parapolybia
9	22	20.2	13	1	NEUT_CHICK	P13724 gallus gall
10	21	19.3	10	1	Q2OG_COMTE	P80466 comamonas t
11	21	19.3	13	1	CXET_CONTE	P81755 conus texti
12	21	19.3	17	1	CHH3_BOMMO	P20729 bombyx mori
13	21	19.3	20	1	CRTC_SPIOL	P30806 spinacia ol
14	20	18.3	12	1	PORD_METTM	P80903 methanobact
15	20	18.3	13	1	HPA1_RANES	P32415 rana escule
16	20	18.3	14	1	CAT2_FASHE	P80342 fasciola he
17	20	18.3	15	1	TA1_TREBR	P34070 tremella br
18	20	18.3	19	1	CAT3_FASHE	P80532 fasciola he
19	20	18.3	19	1	LANA_ACTLG	P56650 actinoplane
20	20	18.3	20	1	CAT1_FASHE	Q09093 fasciola he
21	19.5	17.9	20	1	COG1_PARCM	P20731 paralithode
22	19	17.4	7	1	UH11_RAT	P56576 rattus norv
23	19	17.4	10	1	GON1_PETMA	P04378 petromyzon
24	19	17.4	10	1	GON3_ONCKE	P20367 oncorhynchu
25	19	17.4	10	1	GONL_SQUAC	P27429 squalus aca
26	19	17.4	12	1	RS19_ELYEP	Q47881 elm yellows
27	19	17.4	13	1	FARB_ASCSU	P43173 ascaris suu
28	19	17.4	13	1	NEUT_RANTE	P41536 rana tempor
29	19	17.4	14	1	RS19_CLOPP	Q46228 clover prol
30	19	17.4	14	1	RS19_LOWBP	Q48878 loofah witc
31	19	17.4	15	1	VORA_METTM	P80907 methanobact
32	19	17.4	18	1	NPA_BOVIN	P15506 bos taurus
33	19	17.4	19	1	FIBB_HORSE	P14471 equus cabal
34	19	17.4	19	1	PYRB_PSEFL	P56585 pseudomonas
35	19	17.4	20	1	COG1_CHIOP	P34153 chionoecete
36	18	16.5	8	1	AKH_LIBAU	P25418 libellula a
37	18	16.5	8	1	CCKN_MACEU	P30369 macropus eu
38	18	16.5	8	1	HTF1_PERAM	P04548 periplaneta
39	18	16.5	8	1	HTF2_PERAM	P04549 periplaneta
40	18	16.5	8	1	HTF_TENMO	P25419 tenebrio mo
41	18	16.5	8	1	UF06_MOUSE	P38644 mus musculu
42	18	16.5	10	1	CAER_LITXA	P56264 litoria xan
43	18	16.5	10	1	HTF1_ROMMI	P18110 romalea mic
44	18	16.5	10	1	HTF2_CARMO	P11385 carausius m
45	18	16.5	11	1	BRK_MEGFL	P12797 megascolia

ALIGNMENTS

RESULT 1
TRP3_LEUMA
ID TRP3_LEUMA STANDARD; PRT; 19 AA.
AC P81735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TACHYKININ-RELATED PEPTIDE 3 (LEMTRP 3).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
RN [1]
RP SEQUENCE.
RC TISSUE=MIDGUT;
RX MEDLINE; 97053012.
RA Muren J.E., Naessel D.R.;
RT "Isolation of five tachykinin-related peptides from the midgut of
RT the cockroach Leucophaea madera: existence of N-terminally extended
RT isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 19 19 AMIDATION.
SQ SEQUENCE 19 AA; 1930 MW; 99B5471A011625E5 CRC64;

Query Match 25.7%; Score 28; DB 1; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      7 GEREPG 12
          ||| ||
Db      2 GERAPG 7
```

RESULT 2
AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ADIPOKINETIC HORMONE (AKH) (DIPTERAN CORPORA CARDIACA FACTOR I)
DE (DCC I).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC Tabanidae; Tabanus.
RN [1]
RP SEQUENCE.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE; 90046758.

A;Reference number: A43405; MUID:92388154
A;Accession: A43405
A;Molecule type: protein
A;Residues: 1-10 <VEN>
C;Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phosp
C;Keywords: phosphoric monoester hydrolase; phosphotransferase

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLQGE 8
||:||
Db 5 HLKGE 9

RESULT 14
A60421
hypertrehalosemic hormone - German cockroach
N;Alternate names: Bld-HrTH
C;Species: Blattella germanica (German cockroach)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 31-Oct-1997
C;Accession: A60421; S09137
R;Veenstra, J.A.; Camps, F.
Neuropeptides 15, 107-109, 1990
A;Title: Structure of the hypertrehalosemic neuropeptide of the German cockroach, Blatte
A;Reference number: A60421; MUID:91179584
A;Accession: A60421
A;Molecule type: protein
A;Residues: 1-10 <VEE>
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659
A;Accession: S09137
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 15
S08997
hypertrehalosemic neuropeptide Bld-HrTH - cockroach (Gromphadorina portentosa)
C;Species: Gromphadorina portentosa
C;Date: 30-Jun-1992 #sequence_revision 14-Sep-1994 #text_change 24-Oct-1997
C;Accession: S08997
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659
A;Accession: S08997
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

Search completed: December 21, 2000, 08:30:07
Job time: 271 sec

```
A;Residues: 1-4 <KUR>
C;Keywords: neuropeptide


  Query Match             22.0%;  Score 24;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       2 PGW 4


RESULT   9
A61348
red pigment-concentrating hormone - northern shrimp
N;Alternate names: blanching hormone
C;Species: Pandalus borealis (northern shrimp)
C;Date: 02-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 07-May-1999
C;Accession: A61348; S07139
R;Fernlund, P.; Josefsson, L.
Science 177, 173-175, 1972
A;Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.
A;Reference number: A61348; MUID:72228738
A;Accession: A61348
A;Molecule type: protein
A;Residues: 1-8 <FER1>
R;Fernlund, P.
Biochim. Biophys. Acta 371, 304-311, 1974
A;Title: Structure of the red-pigment-concentrating hormone of the shrimp, Pandalus bore
A;Reference number: S07139; MUID:75054965
A;Accession: S07139
A;Molecule type: protein
A;Residues: 'E',2-8 <FER2>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C;Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pi
zed pigment-containing cells.
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutami
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  10
A33995
adipokinetic hormone - black horse fly
C;Species: Tabanus atratus (black horse fly)
C;Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 31-Oct-1997
C;Accession: A33995
R;Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Nachman, R.J.; Vogel, V.W.; Zhang,
Proc. Natl. Acad. Sci. U.S.A. 86, 8161-8164, 1989
A;Title: Primary structure of two neuropeptide hormones with adipokinetic and hypotrehal
A;Reference number: A33995; MUID:90046758
A;Accession: A33995
A;Molecule type: protein
A;Residues: 1-8 <JAF>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F;8/Modified site: amidated carboxyl end (Trp) #status predicted


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  11
S55310
adipokinetic hormone - damselfly (Pseudagrion inconspicuum)
N;Alternate names: Psi-AKH
C;Species: Pseudagrion inconspicuum
C;Date: 19-Mar-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: S55310
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
A;Accession: S55310
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  12
A58620
adipokinetic hormone - damselfly (Ischnura senegalensis)
C;Species: Ischnura senegalensis
C;Date: 28-Oct-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: A58620
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
A;Accession: A58620
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  13
A43405
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Mar-2000
C;Accession: A43405
R;Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilkis, S.J.; Bartrons, R.
J. Biol. Chem. 267, 17939-17943, 1992
A;Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence
```

```
RESULT   3
S39030
lysyl-bradykinin - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 19-May-1994 #sequence_revision 19-Apr-1996 #text_change 19-Apr-1996
C;Accession: S39030
R;Conlon, J.M.; Olson, K.R.
FEBS Lett. 334, 75-78, 1993
A;Title: Purification of a vasoactive peptide related to lysyl-bradykinin from trout pla
A;Reference number: S39030; MUID:94039817
A;Accession: S39030
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <CON>


  Query Match             26.6%;  Score 29;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 EREPGW 13
           :| |||
Db       1 KRPPGW 6


RESULT   4
PH1788
T cell receptor alpha chain V region (clone 2PBL V alpha 24-4) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1788
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1788
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <POR>


  Query Match             24.8%;  Score 27;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGERE 10
           || |||:
Db       1 HLCGERQ 7


RESULT   5
PH1802
T cell receptor alpha chain V region (clone 3PBL V alpha 24-8) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1802
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>


  Query Match             22.9%;  Score 25;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 8.8e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGER 9
           || |||
Db       1 HLCGER 6


RESULT   6
PH1822
T cell receptor alpha chain V region (clone 5PBL V alpha 24-5) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1822
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1822
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>


  Query Match             22.9%;  Score 25;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 8.8e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGER 9
           || |||
Db       1 HLCGER 6


RESULT   7
F49215
urease (EC 3.5.1.5) small chain UreA - Helicobacter mustelae (ATCC 43772) (fragment)
C;Species: Helicobacter mustelae
C;Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C;Accession: F49215
R;Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A;Title: Purification and characterization of the urease enzymes of Helicobacter spec
A;Reference number: A49215; MUID:93084378
A;Accession: F49215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <TUR>
A;Note: sequence extracted from NCBI backbone (NCBIP:119487)
C;Superfamily: urease 26K chain; urease 11K chain homology; urease 12K chain homology
C;Keywords: hydrolase


  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 9.4e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGE 8
           ::|| ||
Db       9 MMLHYAGE 16


RESULT   8
A34626
RPCH-related neuropeptide - ferruginous spindle
C;Species: Fusinus ferrugineus (ferruginous spindle)
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C;Accession: A34626
R;Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake,
Biochem. Biophys. Res. Commun. 167, 273-279, 1990
A;Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
A;Reference number: A34626; MUID:90179762
A;Accession: A34626
A;Status: preliminary
A;Molecule type: protein
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:07 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-12
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	28.4	18	2	G02018	proteasome chain L
2	30	27.5	16	2	PH1790	T cell receptor al
3	29	26.6	10	2	S39030	lysyl-bradykinin -
4	27	24.8	15	2	PH1788	T cell receptor al
5	25	22.9	17	2	PH1802	T cell receptor al
6	25	22.9	17	2	PH1822	T cell receptor al
7	25	22.9	18	2	F49215	urease (EC 3.5.1.5
8	24	22.0	4	2	A34626	RPCH-related neuro
9	24	22.0	8	2	A61348	red pigment-concen
10	24	22.0	8	2	A33995	adipokinetic hormo
11	24	22.0	8	2	S55310	adipokinetic hormo
12	24	22.0	8	2	A58620	adipokinetic hormo
13	24	22.0	10	2	A43405	6-phosphofructo-2-
14	24	22.0	10	2	A60421	hypertrehalosemic
15	24	22.0	10	2	S08997	hypertrehalosemic
16	24	22.0	10	2	S08998	hypertrehalosemic
17	24	22.0	10	2	A26381	hypertrehalosemic
18	24	22.0	10	2	B33995	hypotrehalosemic h
19	24	22.0	15	2	A26228	spot 42 protein -
20	24	22.0	15	2	PH1329	Ig heavy chain DJ
21	24	22.0	15	2	S38976	chromogranin A - b
22	24	22.0	16	2	PT0282	Ig heavy chain CDR
23	24	22.0	19	2	S29212	protein C - oat (f
24	24	22.0	20	2	A47687	3-dehydroquinate d
25	23	21.1	7	2	A61081	typtophyllin, basi
26	23	21.1	11	2	S68649	spermadhesin AQN-3
27	23	21.1	15	2	S67975	apolipoprotein Cb2
28	23	21.1	16	1	A49761	locustapyrokinin -
29	23	21.1	17	2	PH0778	T-cell receptor al
30	23	21.1	18	2	S71592	serine proteinase
31	23	21.1	19	2	PH1339	Ig heavy chain DJ
32	22	20.2	9	2	S78426	52.5K protein - sp
33	22	20.2	9	2	G56978	collagen alpha 1(I
34	22	20.2	11	2	D56979	collagen alpha 1(I
35	22	20.2	11	2	PD0442	NIPSNAP2 protein -
36	22	20.2	12	2	PQ0776	NADH dehydrogenase
37	22	20.2	13	2	A28505	neurotensin-like p
38	22	20.2	14	2	PH1347	Ig heavy chain DJ
39	22	20.2	14	2	PH1311	Ig heavy chain DJ
40	22	20.2	14	2	PH1321	Ig heavy chain DJ
41	22	20.2	14	2	PH1305	Ig heavy chain DJ
42	22	20.2	14	2	PH1306	Ig heavy chain DJ
43	22	20.2	15	2	PT0222	Ig heavy chain CDR
44	22	20.2	15	2	PH1314	Ig heavy chain DJ
45	22	20.2	15	2	PH1310	Ig heavy chain DJ

ALIGNMENTS

RESULT 1
G02018
proteasome chain LMP7 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 22-Jun-1999
C;Accession: G02018
R;Kim, T.
submitted to the EMBL Data Library, July 1995
A;Reference number: G09054
A;Accession: G02018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <KIM>
A;Cross-references: EMBL:U32862; NID:g1045468; PIDN:AAA80234.1; PID:g1045469
C;Genetics:
A;Gene: LMP7
C;Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 28.4%; Score 31; DB 2; Length 18;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy       9 REPGWIK 15
           :| ||:|
Db       4 KEDGWVK 10
```

RESULT 2
PH1790
T cell receptor alpha chain V region (clone 2PBL V alpha 24-6) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1790
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1790
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16 <POR>

Query Match 27.5%; Score 30; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 HLQGEREP 11
           || ||| |
Db       1 HLCGERGP 8
```

```
Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  14
PCT-US95-09262-13
; Sequence 13, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-13

  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  15
US-08-621-803-227
; Sequence 227, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  227:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.388"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-227


  Query Match             28.4%;  Score 31;  DB 2;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 26;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       1 GWLIQLF 7


Search completed: December 21, 2000, 08:31:49
Job time: 372 sec
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/02465
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noonan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
PCT-US94-02465-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  12
PCT-US95-00498-40
; Sequence 40, Application PC/TUS9500498
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00498
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00498-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  13
PCT-US95-00656-40
; Sequence 40, Application PC/TUS9500656
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Materials and Methods
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00656
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00656-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-13


  Query Match             30.3%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  10
US-09-119-263-40
; Sequence 40, Application US/09119263
; Patent No. 6054431
;  GENERAL INFORMATION:
;    APPLICANT:  Horwitz, Arnold H.
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/119,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,116
;      FILING DATE:
;      APPLICATION NUMBER:  08/372,783
;      FILING DATE:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-09-119-263-40


  Query Match             30.3%;  Score 33;  DB 3;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  11
PCT-US94-02465-40
; Sequence 40, Application PC/TUS9402465
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;       TELEFAX:  312-715-1234
;       TELEX:  910-221-5317
;   INFORMATION FOR SEQ ID NO:  40:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "BPI.38"
US-08-473-344-40


  Query Match           30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   7
US-08-621-803-39
; Sequence 39, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;     APPLICANT:  Better, Marc D.
;     TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;     TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;     NUMBER OF SEQUENCES:  265
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/621,803
;       FILING DATE:  22-MAR-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25,447
;       REFERENCE/DOCKET NUMBER:  27129/33199
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.38"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-39


  Query Match           30.3%;  Score 33;  DB 2;  Length 14;
```

```
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   8
US-08-485-445A-40
; Sequence 40, Application US/08485445A
; Patent No. 5856438
;  GENERAL INFORMATION:
;     APPLICANT:  Little, Roger G.
;     TITLE OF INVENTION:  Biologically Active Peptides from
;     TITLE OF INVENTION:  Functional Domains of Bactericidal/
;     TITLE OF INVENTION:  Permeability-Increasing Protein and
;     TITLE OF INVENTION:  Uses Thereof
;     NUMBER OF SEQUENCES:  226
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;       STREET:  Suite 3400, 500 West Madison Street
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60661
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,445A
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/209,762
;       FILING DATE:  11-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/183,222
;       FILING DATE:  14-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McNicholas, Janet M.
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11018US08/100-224.P4.C1B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-707-8889
;       TELEFAX:  312-707-9155
;   INFORMATION FOR SEQ ID NO:  40:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "BPI.38"
US-08-485-445A-40


  Query Match           30.3%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   9
US-08-621-259A-13
; Sequence 13, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
```

```
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/
;    TITLE OF INVENTION:  Permeability-Increasing Protein and
;    TITLE OF INVENTION:  Uses Thereof
;    NUMBER OF SEQUENCES:  226
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  Suite 3000, 10 S. Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,473A
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-306-473A-40

  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   5
US-08-209-762-40
; Sequence 40, Application US/08209762
; Patent No. 5733872
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,762
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5733872nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-209-762-40

  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   6
US-08-473-344-40
; Sequence 40, Application US/08473344
; Patent No. 5763567
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10  South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,344
;      FILING DATE:  7-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,473
;      FILING DATE:  15-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,762
;      FILING DATE:  11-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-J
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
```

```
US-08-311-611A-40


  Query Match           30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   2
US-08-372-783-40
; Sequence 40, Application US/08372783
; Patent No. 5578572
;  GENERAL INFORMATION:
;    APPLICANT:  Horwitz, Arnold H.
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/372,783
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-08-372-783-40


  Query Match           30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   3
US-08-372-105-40
; Sequence 40, Application US/08372105
; Patent No. 5627153
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Scannon, Patrick J.
;    TITLE OF INVENTION:  Anti-Fungal Materials and Methods
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/372,105
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-08-372-105-40


  Query Match           30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity  71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   4
US-08-306-473A-40
; Sequence 40, Application US/08306473A
; Patent No. 5652332
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:49 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-12
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	30.3	14	1	US-08-311-611A-40	Sequence 40, Appl
2	33	30.3	14	1	US-08-372-783-40	Sequence 40, Appl
3	33	30.3	14	1	US-08-372-105-40	Sequence 40, Appl
4	33	30.3	14	1	US-08-306-473A-40	Sequence 40, Appl
5	33	30.3	14	1	US-08-209-762-40	Sequence 40, Appl
6	33	30.3	14	1	US-08-473-344-40	Sequence 39, Appl
7	33	30.3	14	2	US-08-621-803-39	Sequence 40, Appl
8	33	30.3	14	2	US-08-485-445A-40	Sequence 13, Appl
9	33	30.3	14	2	US-08-621-259A-13	Sequence 40, Appl
10	33	30.3	14	3	US-09-119-263-40	Sequence 40, Appl
11	33	30.3	14	4	PCT-US94-02465-40	Sequence 40, Appl
12	33	30.3	14	4	PCT-US95-00498-40	Sequence 40, Appl
13	33	30.3	14	4	PCT-US95-00656-40	Sequence 13, Appl
14	33	30.3	14	4	PCT-US95-09262-13	Sequence 227, App
15	31	28.4	10	2	US-08-621-803-227	Sequence 219, App
16	31	28.4	10	2	US-08-621-259A-219	Sequence 228, App
17	31	28.4	11	2	US-08-621-803-228	Sequence 220, App
18	31	28.4	11	2	US-08-621-259A-220	Sequence 230, App
19	31	28.4	12	2	US-08-621-803-230	Sequence 222, App
20	31	28.4	12	2	US-08-621-259A-222	Sequence 33, Appl
21	31	28.4	13	1	US-08-261-660A-33	Sequence 42, Appl
22	31	28.4	13	1	US-08-261-660A-42	Sequence 33, Appl
23	31	28.4	13	4	PCT-US94-06931-33	Sequence 42, Appl
24	31	28.4	13	4	PCT-US94-06931-42	Sequence 15, Appl
25	31	28.4	14	1	US-08-311-611A-15	Sequence 33, Appl
26	31	28.4	14	1	US-08-311-611A-33	Sequence 34, Appl
27	31	28.4	14	1	US-08-311-611A-34	Sequence 35, Appl
28	31	28.4	14	1	US-08-311-611A-35	Sequence 36, Appl
29	31	28.4	14	1	US-08-311-611A-36	Sequence 44, Appl
30	31	28.4	14	1	US-08-311-611A-44	Sequence 45, Appl
31	31	28.4	14	1	US-08-311-611A-45	Sequence 46, Appl
32	31	28.4	14	1	US-08-311-611A-46	Sequence 81, Appl
33	31	28.4	14	1	US-08-311-611A-81	Sequence 89, Appl
34	31	28.4	14	1	US-08-311-611A-89	Sequence 91, Appl
35	31	28.4	14	1	US-08-311-611A-91	Sequence 106, App
36	31	28.4	14	1	US-08-311-611A-106	Sequence 107, App
37	31	28.4	14	1	US-08-311-611A-107	Sequence 109, App
38	31	28.4	14	1	US-08-311-611A-109	Sequence 110, App
39	31	28.4	14	1	US-08-311-611A-110	Sequence 122, App
40	31	28.4	14	1	US-08-311-611A-122	Sequence 142, App
41	31	28.4	14	1	US-08-311-611A-142	Sequence 15, Appl
42	31	28.4	14	1	US-08-372-783-15	Sequence 33, Appl
43	31	28.4	14	1	US-08-372-783-33	Sequence 34, Appl
44	31	28.4	14	1	US-08-372-783-34	Sequence 62, Appl
45	31	28.4	14	4	PCT-US95-09262-62	

ALIGNMENTS

```
RESULT 1
US-08-311-611A-40
; Sequence 40, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
;   APPLICANT: Cohen, Jonathan
;   APPLICANT: Kung, Ada H.C.
;   APPLICANT: Lambert, Jr., Lewis H.
;   TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
;   TITLE OF INVENTION: Infection by Administration of
;   TITLE OF INVENTION: Bactericidal/Permeability-Increasing
;   TITLE OF INVENTION:
;   NUMBER OF SEQUENCES: 227
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;     STREET: 6300 Sears Tower, 233 South Wacker Drive
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: USA
;     ZIP: 60606-6402
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/311,611A
;     FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/273,401
;     FILING DATE: 11-JUL-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/125,651
;     FILING DATE: 22-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sharp, Jeffrey S.
;     REGISTRATION NUMBER: 31,879
;     REFERENCE/DOCKET NUMBER: 32251
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312/474-6300
;     TELEFAX: 312/474-0448
;     TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 14 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: misc_feature
;     OTHER INFORMATION: "BPI.38"
```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95105146.
RA Flynn T.R., Hollenberg A.N., Cohen O., Menke J.B., Usala S.J.,
RA Tollin S., Hegarty M.K., Wondisford F.E.;
RT "A novel C-terminal domain in the thyroid hormone receptor selectively
RT mediates thyroid hormone inhibition.";
RL J. Biol. Chem. 269:32713-32716(1994).
SQ SEQUENCE 16 AA; 1920 MW; 79A831A4F3C8E22F CRC64;

Query Match 22.9%; Score 22; DB 4; Length 16;
Best Local Similarity 40.0%; Pred. No. 3.4e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 11 KLFLSLLDFQ 20
|| : : | |
Db 4 KLLMKVTDLQ 13

Search completed: December 21, 2000, 08:35:39
Job time: 602 sec

```
Q9TS18
ID   Q9TS18      PRELIMINARY;      PRT;    20 AA.
AC   Q9TS18;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOTOXIN-BINDING PROTEIN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94039134.
RA   Lutz F., Mohr M., Grimmig M., Leidolf R., Linder D.;
RT   "Pseudomonas aeruginosa cytotoxin-binding protein in rabbit
RT   erythrocyte membranes. An oligomer of 28 kDa with similarity to
RT   transmembrane channel proteins.";
RL   Eur. J. Biochem. 217:1123-1128(1993).
SQ   SEQUENCE   20 AA;  2414 MW;  40B82D7DB5283D2D CRC64;


  Query Match             24.0%;  Score 23;  DB 6;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 KKKLF 13
           |||:|
Db       5 KKKIF 9


RESULT  12
P82138
ID   P82138      PRELIMINARY;      PRT;    20 AA.
AC   P82138;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CHLOROPLAST 30S RIBOSOMAL PROTEIN S15 (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Caryophyllidae; Caryophyllales;
OC   Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. ALWARO; TISSUE=LEAF;
RA   Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT   "Identification of all the proteins in the small subunit of an
RT   organelle (chloroplast) ribosome.";
RL   J. Biol. Chem. 0:0-0(2000).
CC   -!- FUNCTION: THIS PROTEINS BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.
DR   INTERPRO; IPR000589; -.
DR   PROSITE; PS00362; RIBOSOMAL_S15; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2294 MW;  B613012072A648A3 CRC64;


  Query Match             24.0%;  Score 23;  DB 8;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 KKKLFLSLL 17
           ||  |:|::
Db       2 KKNSFISVI 10


RESULT  13
O08806
ID   O08806      PRELIMINARY;      PRT;    20 AA.
AC   O08806;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   SERINE PROTEASE INHIBITOR 14 (SERINE PROTEINASE INHIBITOR MNK26)
DE   (FRAGMENT).
GN   SPI14.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;
RL   J. Biol. Chem. 0:0-0(0).
DR   EMBL; U96709; AAB57821.1; -.
DR   MGD; MGI:894672; Spi14.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2039 MW;  38571040F8263692 CRC64;


  Query Match             24.0%;  Score 23;  DB 11;  Length 20;
  Best Local Similarity   46.7%;  Pred. No. 2.9e+03;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKKLFLSLL 17
           | ||  |  || |::
Db       4 VAASAGKIILFSSMV 18


RESULT  14
P78533
ID   P78533      PRELIMINARY;      PRT;    15 AA.
AC   P78533;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   DEOXYGUANOSINE KINASE (EC 2.7.1.113) (FRAGMENT).
GN   DGK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johansson M., Karlsson A.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + DEOXYGUANOSINE = ADP + DGMP.
DR   EMBL; U62042; AAB48932.1; -.
KW   Transferase.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1706 MW;  53575609CC614F8E CRC64;


  Query Match             22.9%;  Score 22;  DB 4;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 3.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 KLFLSLL 17
           :|||| |
Db       5 RLFLSRL 11


RESULT  15
Q9UD41
ID   Q9UD41      PRELIMINARY;      PRT;    16 AA.
AC   Q9UD41;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   BETA-ISOFORM THYROID HORMONE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
  Query Match             25.0%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 KKLFLSL 16
           |||| :|
Db       2 KKLFAAL 8


RESULT   7
Q9R4E9
ID   Q9R4E9      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4E9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   COA THIOESTERASE (FRAGMENT).
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Rhodopseudomonas.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96171228.
RA   Kuver J., Xu Y., Gibson J.;
RT   "Metabolism of cyclohexane carboxylic acid by the photosynthetic
RT   bacterium Rhodopseudomonas palustris.";
RL   Arch. Microbiol. 164:337-345(1995).
SQ   SEQUENCE   20 AA;  2259 MW;  870D483384D44315 CRC64;

  Query Match             25.0%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 2e+03;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    2;  Gaps    1;

Qy      10 KKLF--LSLLDFQ 20
           | ||  :|:|| :
Db       2 KSLFDLISILDLE 14


RESULT   8
P78199
ID   P78199      PRELIMINARY;      PRT;    17 AA.
AC   P78199;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROBABLE N-ACETYLMURAMOYL-L-ALANINE AMIDASE (EC 3.5.1.28) (FRAGMENT).
GN   AMIA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 97349980.
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).
DR   EMBL; D90874; BAA16324.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   17 AA;  2192 MW;  02EDC1B821D3431D CRC64;

  Query Match             24.0%;  Score 23;  DB 2;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 QASYSKKK 11
            | ::|||:
Db      10 QKAHSKKR 17


RESULT   9
Q56610
ID   Q56610      PRELIMINARY;      PRT;    18 AA.
AC   Q56610;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ACCA (FRAGMENT).
GN   ACCA.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6706;
RX   MEDLINE; 97074686.
RA   Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,
RA   Morris J.G.;
RT   "Cloning and characterization of dnaE, encoding the catalytic subunit
RT   of replicative DNA polymerase III, from Vibrio cholerae strain
RT   C6706.";
RL   Gene 175:281-283(1996).
DR   EMBL; U30472; AAC44579.1; -.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2153 MW;  18EBCDAD212842EF CRC64;

  Query Match             24.0%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 LSLLDFQ 20
           |: |||:
Db       3 LNFLDFE 9


RESULT  10
Q9QVB0
ID   Q9QVB0      PRELIMINARY;      PRT;    18 AA.
AC   Q9QVB0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   (CA2+-MG2+)-ATPASE (FRAGMENT).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93054417.
RA   Ezaki J., Himeno M., Kato K.;
RT   "Purification and characterization of (Ca2+-Mg2+)-ATPase in rat liver
RT   lysosomal membranes.";
RL   J. Biochem. 112:33-39(1992).
SQ   SEQUENCE   18 AA;  2008 MW;  8314BA373ED6AA47 CRC64;

  Query Match             24.0%;  Score 23;  DB 11;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 KKKLFLSLL 17
            |:::|: ||
Db       6 KERVFMVLL 14


RESULT  11
```

```
GN   NISG.
OS   Lactococcus lactis.
OG   Plasmid pLEB513.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Wahlstroem G., Takala T., Saris P.E.J.;
RT   "Evidence for a mosaic structure of the Tn5481 in Lactococcus lactis
RT   N8.";
RL   DNA Seq. 9:245-261(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Saris P.E.J.;
RT   "Characterization of the nisFEG operon of the nisin Z producing
RT   Lactococcus lactis subsp. lactis N8 strain.";
RL   DNA Seq. 9:263-274(1998).
DR   EMBL; AJ000993; CAA04441.1; -.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2139 MW;  6A429DC80E673613 CRC64;


  Query Match             27.1%;  Score 26;  DB 2;  Length 19;
  Best Local Similarity   41.7%;  Pred. No. 8.6e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSKKKL 12
          | :  : ||||:
Db      7 TIILVALSKKKI 18


RESULT   3
P97522
ID   P97522      PRELIMINARY;      PRT;    18 AA.
AC   P97522;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   CFTR GENE.
GN   CFTR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Vuillaumier S., Denamur E.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X95927; CAA65168.1; -.
SQ   SEQUENCE   18 AA;  2158 MW;  5C5855056C1CE6DE CRC64;


  Query Match             26.0%;  Score 25;  DB 11;  Length 18;
  Best Local Similarity   38.9%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDFQ 20
          :| |  :|  |:| | |:
Db      1 MQKSPLEKASFISKLFFR 18


RESULT   4
Q9URE0
ID   Q9URE0      PRELIMINARY;      PRT;    15 AA.
AC   Q9URE0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   TYPE II TOPOISOMERASE, TOPOISOMERASE II.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Elsea S.H., Hsiung Y., Nitiss J.L., Osheroff N.;
RL   J. Biol. Chem. 270:1913-1920(1995).
SQ   SEQUENCE   15 AA;  2094 MW;  0A6A37F6E81E85F6 CRC64;


  Query Match             25.0%;  Score 24;  DB 3;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      7 YSKKKLFLS 15
          | |:| ::|
Db      2 YQKRKDYMS 10


RESULT   5
Q9UCH4
ID   Q9UCH4      PRELIMINARY;      PRT;    15 AA.
AC   Q9UCH4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NATURAL KILLER ENHANCING FACTOR, NKEF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93215001.
RA   Shau H., Gupta R.K., Golub S.H.;
RT   "Identification of a natural killer enhancing factor (NKEF) from human
RT   erythroid cells.";
RL   Cell. Immunol. 147:1-11(1993).
SQ   SEQUENCE   15 AA;  1738 MW;  4909D4A793D382BF CRC64;


  Query Match             25.0%;  Score 24;  DB 4;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFL 14
          |||   |  :||
Db      2 VQAFQGKVNVFL 13


RESULT   6
O06514
ID   O06514      PRELIMINARY;      PRT;    18 AA.
AC   O06514;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
DE   DNA, MOSAIC MERCURY RESISTANCE TRANSPOSABLE ELEMENT
DE   (MER-OPERON)1040 BP (FRAGMENT).
GN   MERP.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CH210; TRANSPOSON=TN5059;
RX   MEDLINE; 97303088.
RA   Nikiforov V., Kholodii G., Minakhin L., Gorlenko Z., Kalyaeva E.,
RA   Mindlin S., Nikiforov V.;
RT   "Intercontinental spread of promiscuous mercury-resistance transposons
RT   in environmental bacteria.";
RL   Mol. Microbiol. 24:321-329(1997).
DR   EMBL; Y09026; CAA70241.1; -.
FT   NON_TER       18     18
SQ   SEQUENCE   18 AA;  1899 MW;  484A2D7197814DF9 CRC64;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:39 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-11
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	27.1	18	3	Q02414	Q02414 aspergillus
2	26	27.1	19	2	Q9RLU4	Q9rlu4 lactococcus
3	25	26.0	18	11	P97522	P97522 rattus norv
4	24	25.0	15	3	Q9URE0	Q9ure0 saccharomyc
5	24	25.0	15	4	Q9UCH4	Q9uch4 homo sapien
6	24	25.0	18	2	O06514	O06514 escherichia
7	24	25.0	20	2	Q9R4E9	Q9r4e9 rhodopseudo
8	23	24.0	17	2	P78199	P78199 escherichia
9	23	24.0	18	2	Q56610	Q56610 vibrio chol
10	23	24.0	18	11	Q9QVB0	Q9qvb0 rattus sp.
11	23	24.0	20	6	Q9TS18	Q9ts18 oryctolagus
12	23	24.0	20	8	P82138	P82138 spinacia ol
13	23	24.0	20	11	O08806	O08806 mus musculu
14	22	22.9	15	4	P78533	P78533 homo sapien
15	22	22.9	16	4	Q9UD41	Q9ud41 homo sapien
16	22	22.9	16	6	Q9TRA2	Q9tra2 sus scrofa
17	22	22.9	17	11	Q9Z170	Q9z170 mus musculu
18	22	22.9	19	4	Q9UMM9	Q9umm9 homo sapien
19	22	22.9	19	6	Q9TRF3	Q9trf3 macropus gi
20	21	21.9	9	2	Q45852	Q45852 clostridium
21	21	21.9	12	11	O54970	O54970 mus musculu
22	21	21.9	14	2	Q9R5Q6	Q9r5q6 burkholderi
23	21	21.9	18	11	Q9QUZ8	Q9quz8 rattus sp.
24	21	21.9	19	3	Q92321	Q92321 saccharomyc
25	21	21.9	20	2	Q49448	Q49448 mycoplasma
26	21	21.9	20	9	Q37868	Q37868 bacteriopha
27	21	21.9	20	12	Q86935	Q86935 human herpe
28	21	21.9	20	12	Q86936	Q86936 human herpe
29	21	21.9	20	12	Q86937	Q86937 human herpe
30	21	21.9	20	12	Q86938	Q86938 human herpe
31	21	21.9	20	12	Q86939	Q86939 human herpe
32	21	21.9	20	12	Q86940	Q86940 human herpe
33	21	21.9	20	12	Q86941	Q86941 human herpe
34	21	21.9	20	12	Q86942	Q86942 human herpe
35	20.5	21.4	19	13	Q9PRP1	Q9prp1 gallus gall
36	20	20.8	13	11	Q62354	Q62354 mus musculu
37	20	20.8	13	11	Q62355	Q62355 mus musculu
38	20	20.8	15	2	O68425	O68425 buchnera ap
39	20	20.8	15	2	Q9R4Y0	Q9r4y0 pseudomonas
40	20	20.8	16	2	Q47605	Q47605 escherichia
41	20	20.8	16	2	Q9R4L0	Q9r4l0 spiroplasma
42	20	20.8	16	10	Q9S8A4	Q9s8a4 secale cere
43	20	20.8	16	12	Q79458	Q79458 human immun
44	20	20.8	17	8	Q9ZY77	Q9zy77 trissolcus
45	20	20.8	18	2	Q46390	Q46390 clostridium

ALIGNMENTS

RESULT 1
Q02414
ID Q02414 PRELIMINARY; PRT; 18 AA.
AC Q02414;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE OROTIDINE-5'-MONOPHOSPHATE DECARBOXYLASE (FRAGMENT).
GN PYRG.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
OC anamorphic Trichocomaceae; Aspergillus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96031709.
RA Gouka R.J., Hessing J.G., Stam H., Musters W., Hondel C.A.;
RT "A novel strategy for the isolation of defined pyrG mutants and the
RT development of a site-specific integration system for Aspergillus
RT awamori.";
RL Curr. Genet. 27:536-540(1995).
DR EMBL; S79674; AAB35350.1; -.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 1891 MW; E1A0E34C7D23688E CRC64;

Query Match 27.1%; Score 26; DB 3; Length 18;
Best Local Similarity 41.2%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

Qy 2 TVQASYSKKKLFLSLLD 18
|| | : | | | | |
Db 2 TVSADVTTTKELLDLAD 18

RESULT 2
Q9RLU4
ID Q9RLU4 PRELIMINARY; PRT; 19 AA.
AC Q9RLU4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE NISG PROTEIN (FRAGMENT).

```
RL   Arch. Insect Biochem. Physiol. 39:1-8(1998).
CC   -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC       AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC       BIOSYNTHESIS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC       EF-TU/EF-1A SUBFAMILY.
DR   INTERPRO; IPR000795; -.
DR   PROSITE; PS00301; EFACTOR_GTP; PARTIAL.
KW   Elongation factor; Protein biosynthesis; GTP-binding.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1670 MW;  1EB7DA05B09B3751 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 15;
  Best Local Similarity   20.0%;  Pred. No. 2.3e+03;
  Matches    2;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 SKKKLFLSLL 17
           :|:|: ::::
Db       1 AKEKIHINIV 10


RESULT  14
RL6_VIBPR
ID   RL6_VIBPR      STANDARD;      PRT;    16 AA.
AC   Q56715;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   50S RIBOSOMAL PROTEIN L6 (FRAGMENT).
GN   RPLF.
OS   Vibrio proteolyticus (Aeromonas proteolytica).
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15338;
RX   MEDLINE; 97149305.
RA   Setterquist R.A., Smith G.K., Oakley T.H., Lee Y.H., Fox G.E.;
RT   "Sequence, overproduction and purification of Vibrio proteolyticus
RT   ribosomal protein L18 for in vitro and in vivo studies.";
RL   Gene 183:237-242(1996).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC       LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE
CC       PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U38943; AAB41328.1; -.
DR   INTERPRO; IPR002358; -.
DR   PROSITE; PS00525; RIBOSOMAL_L6_1; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER      1      1
SQ   SEQUENCE   16 AA;  1935 MW;  ABC19078DF581B6C CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           |:   :|||
Db       8 VRTKEAKKK 16


RESULT  15
PHSL_DESBN
ID   PHSL_DESBN     STANDARD;      PRT;    19 AA.
AC   P13066;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PERIPLASMIC [NIFESE] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESE
DE   HYDROGENLYASE LARGE CHAIN) (FRAGMENT).
OS   Desulfovibrio baculatus (strain Norway 4).
OC   Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88106446.
RA   Prickril B.C., He S.H., Li C., Menon N., Choi E.S., Przybyla A.E.,
RA   Dervartanian D.V., Peck H.D. Jr., Fauque G., le Gall J., Teixeira M.,
RA   Moura I., Moura J.J.G., Patil D., Huynh B.H.;
RT   "Identification of three classes of hydrogenase in the genus,
RT   Desulfovibrio.";
RL   Biochem. Biophys. Res. Commun. 149:369-377(1987).
CC   -!- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) = 2 OXIDIZED
CC       FERREDOXIN + H(2).
CC   -!- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE
CC       SELENOCYSTEINE.
CC   -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE
CC       SUBUNIT FAMILY.
DR   PIR; H27480; H27480.
DR   INTERPRO; IPR001501; -.
DR   PROSITE; PS00507; NI_HGENASE_L_1; PARTIAL.
DR   PROSITE; PS00508; NI_HGENASE_L_2; PARTIAL.
KW   Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1942 MW;  2BFCD2D360F00367 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ASYSKKKL 12
           |:  |||:
Db       7 AADGKKKI 14


Search completed: December 21, 2000, 08:32:59
Job time: 442 sec
```

```
FT   NON_TER       15     15
SQ   SEQUENCE   15 AA;  1732 MW;  53CC9A3CADDDDAA1 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 SYSKKKL 12
           :| | ||
Db       1 NYMKPKL 7


RESULT  10
RIPK_TRIKI
ID   RIPK_TRIKI     STANDARD;      PRT;     16 AA.
AC   P16093;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   RIBOSOME-INACTIVATING PROTEIN TRICHOKIRIN (RRNA N-GLYCOSIDASE)
DE   (EC 3.2.2.22) (FRAGMENT).
OS   Trichosanthes kirilowii (Mongolian snake-gourd).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Cucurbitales; Cucurbitaceae; Trichosanthes.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SEED;
RX   MEDLINE; 89005108.
RA   Casellas P., Dussossoy D., Falasca A.I., Barbieri L.,
RA   Guillemot J.C., Ferrara P., Bolognesi A., Cenini P., Stirpe F.;
RT   "Trichokirin, a ribosome-inactivating protein from the seeds of
RT   Trichosanthes kirilowii Maximowicz. Purification, partial
RT   characterization and use for preparation of immunotoxins.";
RL   Eur. J. Biochem. 176:581-588(1988).
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC       SPECIFIC ADENOSINE ON THE 28S RRNA.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING
CC       PROTEINS. BELONGS TO TYPE 1 RIP.
DR   PIR; S01669; S01669.
DR   INTERPRO; IPR001574; -.
DR   PROSITE; PS00275; SHIGA_RICIN; PARTIAL.
KW   Protein synthesis inhibitor; Hydrolase; Toxin; Glycoprotein.
FT   NON_TER       16     16
SQ   SEQUENCE   16 AA;  1605 MW;  5E268A7F345935A2 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ASYSK 9
           ||| |
Db      12 ASYEK 16


RESULT  11
B44K_PORGI
ID   B44K_PORGI     STANDARD;      PRT;      8 AA.
AC   P81886;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   44 KDA IMMUNOGENIC PROTEIN (FRAGMENT).
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; CFB group; Bacteroidaceae; Porphyromonas.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VPB 3492;
RX   MEDLINE; 20198497.
RA   Norris J.M., Love D.N.;
RT   "Serum antibody responses of cats to soluble whole cell antigens of
RT   feline porphyromonas gingivalis.";
RL   Vet. Microbiol. 73:37-49(2000).
CC   -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
KW   Antigen.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  989 MW;  9554540326CB476D CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 ASYSKKKL 12
             | | |: :
Db       1 APYQKRNI 8


RESULT  12
MAST_VESBA
ID   MAST_VESBA     STANDARD;      PRT;     14 AA.
AC   P21654;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   MASTOPARAN B.
OS   Vespa basalis (Hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 91174755.
RA   Lo C.-L., Hwang L.-L.;
RT   "Structure and biological activities of a new mastoparan isolated
RT   from the venom of the hornet Vespa basalis.";
RL   Biochem. J. 274:453-456(1991).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC       THAT COUPLE TO PHOSPHOLIPASE C.
DR   PIR; S14336; S14336.
KW   Mast cell degranulation; Venom; Amidation.
FT   MOD_RES       14     14       AMIDATION.
SQ   SEQUENCE   14 AA;  1613 MW;  D35944CA193A19A2 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 SYSKKKL 12
           |::|| |
Db       8 SWAKKVL 14


RESULT  13
EF1A_MICCR
ID   EF1A_MICCR     STANDARD;      PRT;     15 AA.
AC   P81266;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ELONGATION FACTOR 1-ALPHA (EF-1-ALPHA) (FRAGMENT).
OS   Microplitis croceipes.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC   Ichneumonoidea; Braconidae; Microgastrinae; Microplitis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 99033459.
RA   Stuart M.K.;
RT   "An antibody diagnostic for hymenopteran parasitism is specific for a
RT   homologue of elongation factor-1 alpha.";
```

```
RESULT   6
MI17_BOVIN
ID   MI17_BOVIN     STANDARD;      PRT;    20 AA.
AC   P35451;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   17 KDA MILK GLYCOPROTEIN (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=MILK;
RX   MEDLINE; 93308294.
RA   Soerensen E.S., Petersen T.E.;
RT   "Purification and characterization of three proteins isolated from
RT   the proteose peptone fraction of bovine milk.";
RL   J. Dairy Res. 60:189-197(1993).
CC   -!- PTM: N-GLYCOSYLATED.
CC   -!- SIMILARITY: TO CAMEL WHEY PROTEIN.
KW   Glycoprotein; Milk.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2233 MW;  4CCA589404C62C27 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       1 TTVQASYSKKKLFLSLL 17
           :: |      || ||:|
Db       1 SSXQPQSQNPKLPLSIL 17


RESULT   7
TL18_SPIOL
ID   TL18_SPIOL     STANDARD;      PRT;    20 AA.
AC   P82536;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   THYLAKOID LUMENAL 18 KDA PROTEIN (P18) (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RA   Kieselbach T., Bystedt M., Schroeder W.P.;
RL   Submitted (MAY-2000) to the SWISS-PROT data bank.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.
KW   Chloroplast; Thylakoid membrane.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2192 MW;  752C21963F49FA64 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TTVQASYSKKKLF 13
           | :|:  : | :|
Db       5 TPLQSKVTNKVVF 17


RESULT   8
PA2B_VIPBO
ID   PA2B_VIPBO     STANDARD;      PRT;    12 AA.
AC   P31859;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2, BASIC (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE   2-ACYLHYDROLASE) (FRAGMENT).
OS   Vipera berus orientalis (Viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Vipera.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 74128698.
RA   Delori P.J.;
RT   "Purification and physiochemical, chemical and biological properties
RT   of a toxic A2 phospholipase isolated from the venom of viperidae
RT   snakes: Vipera berus.";
RL   Biochimie 55:1031-1045(1973).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   INTERPRO; IPR001211; -.
DR   PROSITE; PS00118; PA2_HIS; PARTIAL.
DR   PROSITE; PS00119; PA2_ASP; PARTIAL.
KW   Hydrolase; Lipid degradation; Calcium; Venom.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1412 MW;  D054351A89969879 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 SLLDF 19
           |||:|
Db       1 SLLEF 5


RESULT   9
FGF1_CANFA
ID   FGF1_CANFA     STANDARD;      PRT;    15 AA.
AC   P18651;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HEPARIN-BINDING GROWTH FACTOR 1 (HBGF-1) (ACIDIC FIBROBLAST GROWTH
DE   FACTOR) (AFGF) (ALPHA-ENDOTHELIAL CELL GROWTH FACTOR) (FRAGMENT).
GN   FGF1 OR FGF-1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89231704.
RA   Quinkler W., Maasberg M., Bernotat-Danielowski S., Luethe N.,
RA   Sharma H.S., Schaper W.;
RT   "Isolation of heparin-binding growth factors from bovine, porcine and
RT   canine hearts.";
RL   Eur. J. Biochem. 181:67-73(1989).
CC   -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS
CC       IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN
CC       VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND
CC       CONCENTRATION OF THESE 2 GROWTH FACTORS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY
CC       THAN DOES BFGF.
CC   -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
DR   PIR; S03955; S03955.
DR   HSSP; P05230; 2AXM.
DR   INTERPRO; IPR002209; -.
DR   PROSITE; PS00247; HBGF_FGF; PARTIAL.
KW   Growth factor; Mitogen; Vascularization; Heparin-binding.
```

```
RX   MEDLINE; 79255521.
RA   Possani L.D., Alagon A.C., Fletcher P.L. Jr., Varela M.J., Julia J.Z.;
RT   "Purification and characterization of a phospholipase A2 from the
RT   venom of the coral snake, Micrurus fulvius microgalbineus (Brown and
RT   Smith).";
RL   Biochem. J. 179:603-606(1979).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   INTERPRO; IPR001211; -.
DR   PROSITE; PS00118; PA2_HIS; PARTIAL.
DR   PROSITE; PS00119; PA2_ASP; PARTIAL.
KW   Hydrolase; Lipid degradation; Calcium; Venom.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1398 MW;  CC21992A899F0339 CRC64;


  Query Match             24.0%;  Score 23;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 SLLDFQ 20
           |||:|:
Db       1 SLLBFK 6


RESULT   3
CPBX_CAVPO
ID   CPBX_CAVPO     STANDARD;      PRT;     20 AA.
AC   P34033;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CYTOCHROME P450IIB (EC 1.14.14.1) (FRAGMENT).
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN   [1]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   STRAIN=HARTLEY; TISSUE=LIVER;
RX   MEDLINE; 91054472.
RA   Narimatsu S., Akutsu Y., Matsunaga T., Watanabe K., Yamamoto I.,
RA   Yoshimura H.;
RT   "Purification of a cytochrome P450 isozyme belonging to a subfamily
RT   of P450 IIB from liver microsomes of guinea pigs.";
RL   Biochem. Biophys. Res. Commun. 172:607-613(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME IS ACTIVE
CC       UPON P.NITROANISOLE, ANILINE, D-BENZPHETAMINE, DELTA(9)-
CC       TETRAHYDROCANNABINOL (THC) AND STRYCHNINE.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A36154; A36154.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2259 MW;  78DC81280C970A55 CRC64;


  Query Match             24.0%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LFLSLL 17
           |||:||
Db       7 LFLALL 12


RESULT   4
VR90_BORPE
ID   VR90_BORPE     STANDARD;      PRT;     20 AA.
AC   P81549;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   VIRULENCE-ASSOCIATED OUTER MEMBRANE PROTEIN VIR90 (FRAGMENT).
GN   VIR90.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC   Bordetella.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=TOHAMA I;
RX   MEDLINE; 99179239.
RA   Passerini de Rossi B.N., Friedman L.E., Gonzalez Flecha F.L.,
RA   Castello P.R., Franco M.A., Rossi J.P.F.C.;
RT   "Identification of Bordetella pertussis virulence-associated outer
RT   membrane proteins.";
RL   FEMS Microbiol. Lett. 172:9-13(1999).
KW   Outer membrane; Virulence.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2116 MW;  D24E1CDCA665206C CRC64;


  Query Match             22.9%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 9.4e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TVQASYS 8
           ||:  ||
Db      11 TVEGEYS 17


RESULT   5
XYNB_DICB4
ID   XYNB_DICB4     STANDARD;      PRT;     10 AA.
AC   P80717;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ENDO-1,4-BETA-XYLANASE B (EC 3.2.1.8) (XYLANASE B)
DE   (1,4-BETA-D-XYLAN XYLANOHYDROLASE B) (FRAGMENT).
OS   Dictyoglomus sp. (strain B4A).
OC   Bacteria; Dictyoglomus group; Dictyoglomus.
RN   [1]
RP   SEQUENCE.
RA   Adamsen A.K., Jacobsen S., Ahring B.K.;
RL   Submitted (OCT-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC       LINKAGES IN XYLANS.
CC   -!- PATHWAY: XYLAN DEGRADATION.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC       HYDROLASES).
DR   INTERPRO; IPR001000; -.
DR   PROSITE; PS00591; GLYCOSYL_HYDROL_F10; PARTIAL.
KW   Xylan degradation; Hydrolase; Glycosidase.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1144 MW;  4554322AA72041A3 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   55.6%;  Pred. No. 6.8e+02;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      10 KKLFLSLLD 18
           ||  | | |
Db       2 KKTILDLKD 10
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:59 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-11
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	24	25.0	18	1	UC21_MAIZE	P80627	zea mays (m
2	23	24.0	12	1	PA21_MICFM	P25072	micrurus fu
3	23	24.0	20	1	CPBX_CAVPO	P34033	cavia porce
4	22	22.9	20	1	VR90_BORPE	P81549	bordetella
5	21	21.9	10	1	XYNB_DICB4	P80717	dictyoglomu
6	21	21.9	20	1	MI17_BOVIN	P35451	bos taurus
7	21	21.9	20	1	TL18_SPIOL	P82536	spinacia ol
8	20	20.8	12	1	PA2B_VIPBO	P31859	vipera beru
9	20	20.8	15	1	FGF1_CANFA	P18651	canis famil
10	20	20.8	16	1	RIPK_TRIKI	P16093	trichosanth
11	19	19.8	8	1	B44K_PORGI	P81886	porphyromon
12	19	19.8	14	1	MAST_VESBA	P21654	vespa basal
13	19	19.8	15	1	EF1A_MICCR	P81266	microplitis
14	19	19.8	16	1	RL6_VIBPR	Q56715	vibrio prot
15	19	19.8	19	1	PHSL_DESBN	P13066	desulfovibr
16	19	19.8	20	1	STYA_STYCL	P81469	styela clav
17	18	18.8	9	1	ULAE_HUMAN	P31931	homo sapien
18	18	18.8	13	1	UVRD_SALTY	Q05311	salmonella
19	18	18.8	13	1	YPE2_LACLC	P42021	lactococcus
20	18	18.8	16	1	ARCD_PSEPU	P41147	pseudomonas
21	18	18.8	17	1	ATPI_PAVLU	P28529	pavlova lut
22	18	18.8	19	1	UP24_UPEIN	P82030	uperoleia i
23	18	18.8	20	1	AMP_FUSNU	P81207	fusobacteri
24	18	18.8	20	1	CRP_MUSCA	P19094	mustelus ca
25	17	17.7	8	1	CCKN_MACEU	P30369	macropus eu
26	17	17.7	11	1	TKN_ELEMO	P01293	eledone mos
27	17	17.7	12	1	NO40_SOYBN	P55960	glycine max
28	17	17.7	14	1	MAST_VESMA	P04205	vespa manda
29	17	17.7	14	1	MAST_VESXA	P01515	vespa xanth
30	17	17.7	14	1	UC34_MAIZE	P80640	zea mays (m
31	17	17.7	15	1	CBPB_PROAT	P19628	protopterus
32	17	17.7	17	1	BOL1_MEGPE	P10521	megabombus
33	17	17.7	18	1	NODD_RHILO	Q52838	rhizobium l
34	17	17.7	18	1	RL23_HALCU	P05975	halobacteri
35	17	17.7	20	1	CATA_ACIRA	P81422	acinetobact
36	17	17.7	20	1	RIPX_CUCPE	P80750	cucurbita p
37	16	16.7	8	1	AKH_MELML	P25423	melolontha
38	16	16.7	8	1	COXG_RAT	P80430	rattus norv
39	16	16.7	8	1	CPD1_ENTFA	P13269	enterococcu
40	16	16.7	10	1	COXM_RAT	P80431	rattus norv
41	16	16.7	10	1	PNEU_RAT	P21996	rattus norv
42	16	16.7	11	1	CSI5_BACSU	P81095	bacillus su
43	16	16.7	11	1	UXB2_YEAST	P99013	saccharomyc
44	16	16.7	14	1	MAST_VESLE	P01514	vespula lew
45	16	16.7	14	1	RS19_PPWBP	Q52093	pigeon pea

ALIGNMENTS

RESULT 1
UC21_MAIZE
ID UC21_MAIZE STANDARD; PRT; 18 AA.
AC P80627;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 443)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE.
RC TISSUE=COLEOPTILE;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.9, ITS MW IS: 26.5 KDA.
CC -!- SIMILARITY: TO L-ASCORBATE PEROXIDASES.
DR HSSP; P48534; 1APX.
DR MAIZE-2DPAGE; P80627; COLEOPTILE.
DR MAIZEDB; 123953; -.
FT NON_TER 1 1
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1938 MW; F32F6FEF038BAB8A CRC64;

Query Match 25.0%; Score 24; DB 1; Length 18;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 TVQASYS 8
          || | ||
Db      6 TVSAEYS 12
```

RESULT 2
PA21_MICFM
ID PA21_MICFM STANDARD; PRT; 12 AA.
AC P25072;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE 2-ACYLHYDROLASE) (FRAGMENT).
OS Micrurus fulvius microgalbineus (Mexican coral snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Micrurus.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;

A;Title: Purification of the sequence-specific transcription factor CTCBF, involved in t
A;Reference number: S54272; MUID:95188883
A;Accession: S54272
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <GEN>

Query Match 22.9%; Score 22; DB 2; Length 18;
Best Local Similarity 33.3%; Pred. No. 2.6e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSKKKL 12
          :| :  ||:::|
Db      6 STGKVEYSEEEL 17
```

RESULT 14
G22565
R-phycoerythrin gamma-B chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: G22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: G22565
A;Molecule type: protein
A;Residues: 1-13 <KLO>

Query Match 21.9%; Score 21; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSK 9
          |  ||:: |
Db      5 TVPQAAFEK 13
```

RESULT 15
S29209
avenin alpha-2 - oat (fragment)
N;Alternate names: CIP-3; coeliac immunoreactive protein 3
C;Species: Avena sativa (oat)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C;Accession: S29209
R;Rocher, A.; Colilla, F.; Ortiz, M.L.; Mendez, E.
FEBS Lett. 310, 37-40, 1992
A;Title: Identification of the three major coeliac immunoreactive proteins and one alpha
A;Reference number: S29207; MUID:92405739
A;Accession: S29209
A;Molecule type: protein
A;Residues: 1-14 <ROC>
A;Experimental source: endosperm
C;Superfamily: gliadin
C;Keywords: prolamin; seed

Query Match 21.9%; Score 21; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSKK 10
          |||| : |::
Db      2 TTVQYNPSEQ 11
```

Search completed: December 21, 2000, 08:30:07
Job time: 271 sec

```
RESULT   8
S39049
cytotoxin-binding protein - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 16-Feb-1997
C;Accession: S39049
R;Lutz, F.; Mohr, M.; Grimmig, M.; Leidolf, R.; Linder, D.
Eur. J. Biochem. 217, 1123-1128, 1993
A;Title: Pseudomonas aeruginosa cytotoxin-binding protein in rabbit erythrocyte membrane
A;Reference number: S39049; MUID:94039134
A;Accession: S39049
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <LUT>
C;Superfamily: lens fiber membrane major intrinsic protein

  Query Match             24.0%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 KKKLF 13
            |||:|
Db       5 KKKIF 9


RESULT   9
PH0786
T-cell receptor alpha chain (I7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0786
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0786
A;Molecule type: mRNA
A;Residues: 1-13 <CAS>
A;Cross-references: EMBL:X60889
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             22.9%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   33.3%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ASYSKKKLF 13
            ::|: | :|
Db       4 SAYANKMIF 12


RESULT  10
S65717
prostaglandin D-synthase - rat (fragment)
N;Alternate names: prostaglandin-H2 D-isomerase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 31-Oct-1997
C;Accession: S65717
R;Giacomelli, S.; Leone, M.G.; Grima, J.; Silvestrini, B.; Cheng, C.Y.
Biochim. Biophys. Acta 1310, 269-276, 1996
A;Title: Astrocytes synthesize and secrete prostaglandin D synthetase in vitro.
A;Reference number: S65716; MUID:96177373
A;Accession: S65717
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <GIA>
C;Superfamily: lipocalin; lipocalin homology

  Query Match             22.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TVQASYSKKK 11
            ||| :: : |
Db       3 TVQPNFQQDK 12


RESULT  11
PA0046
protein QA100044 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997
C;Accession: PA0046; PA0042
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensi
A;Reference number: PA0001
A;Accession: PA0046
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: stem

  Query Match             22.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 KKLFLSLLD 18
            || ||:: |
Db       4 KKGFLAVSD 12


RESULT  12
S34444
blaZ protein - Staphylococcus aureus plasmid pI258 (fragment)
C;Species: Staphylococcus aureus
C;Date: 22-Nov-1993 #sequence_revision 09-Mar-1996 #text_change 07-May-1999
C;Accession: S34444
R;Wang, P.Z.; Projan, S.J.; Novick, R.P.
Nucleic Acids Res. 19, 4000, 1991
A;Title: Nucleotide sequence of beta-lactamase regulatory genes from staphylococcal p
A;Reference number: S34444; MUID:91319567
A;Accession: S34444
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-16 <WAN>
A;Cross-references: EMBL:M62650
A;Experimental source: strain RN11
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
C;Genetics:
A;Gene: blaZ
A;Genome: plasmid pI258

  Query Match             22.9%;  Score 22;  DB 2;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2.4e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 KKLFLSLLDF 19
            |:| | :| |
Db       6 KQLQLKILTF 15


RESULT  13
S54272
CTC 75 protein - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S54272
R;Genersch, E.; Eckerskorn, C.; Lottspeich, F.; Herzog, C.; Kuehn, K.; Poeschl, E.
EMBO J. 14, 791-800, 1995
```

```
          :|||||  |
Db      3 EKKLFLKAL 11


RESULT   3
I52618
hemoglobin beta chain thalassemia mutant Portuguese - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 11-Jul-1996 #text_change 20-Apr-2000
C;Accession: I52618
R;Oner, R.; Oner, C.; Wilson, J.B.; Tamagnini, G.P.; Ribeiro, L.M.; Huisman, T.H.
Br. J. Haematol. 79, 306-310, 1991
A;Title: Dominant beta-thalassaemia trait in a Portuguese family is caused by a deletion
A;Reference number: I52618; MUID:92068764
A;Accession: I52618
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <ONE>
A;Cross-references: GB:S68042; NID:g239717; PIDN:AAB20440.1; PID:g239718
C;Genetics:
A;Gene: GDB:HBB
A;Cross-references: GDB:119297; OMIM:141900
A;Map position: 11p15.4-11p15.4


  Query Match             27.1%;  Score 26;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 4.5e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 VQASYSK 9
          |||:| |
Db      3 VQAAYQK 9


RESULT   4
A61411
ameletin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
C;Accession: A61411
R;Burzynski, S.R.
Anal. Biochem. 70, 359-365, 1976
A;Title: Sequential analysis in subnanomolar amounts of peptides. Determination of the s
A;Reference number: A61411; MUID:76182447
A;Accession: A61411
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6 <BUR>
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental


  Query Match             26.0%;  Score 25;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 QASYSK 9
          || |||
Db      1 QAGYSK 6


RESULT   5
I49593
cystic fibrosis transmembrane conductance regulator - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I49593
R;Denamur, E.; Chehab, F.F.
Hum. Mol. Genet. 3, 1089-1094, 1994
A;Title: Analysis of the mouse and rat CFTR promoter regions.
A;Reference number: I49593; MUID:95072572
A;Accession: I49593
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-17 <RES>
A;Cross-references: GB:L04873; NID:g414726; PIDN:AAA73562.1; PID:g553892
C;Genetics:
A;Gene: CFTR
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding casse


  Query Match             25.0%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   41.2%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDF 19
          :| |  :|  |:| | |
Db      1 MQKSPLEKASFISKLFF 17


RESULT   6
I84733
gene CFTR protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I84733
R;Denamur, E.; Chehab, F.F.
Hum. Mol. Genet. 3, 1089-1094, 1994
A;Title: Analysis of the mouse and rat CFTR promoter regions.
A;Reference number: I49593; MUID:95072572
A;Accession: I84733
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-17 <RES>
A;Cross-references: GB:L26098; NID:g425185; PIDN:AAA73561.1; PID:g915270
C;Genetics:
A;Gene: CFTR
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding casse


  Query Match             25.0%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   41.2%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDF 19
          :| |  :|  |:| | |
Db      1 MQKSPLEKASFISKLFF 17


RESULT   7
A49192
transthyretin - bullfrog (fragment)
N;Alternate names: 3,5,3'-L-triiodothyronine-specific binding protein
C;Species: Rana catesbeiana (bullfrog)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C;Accession: A49192
R;Yamauchi, K.; Kasahara, T.; Hayashi, H.; Horiuchi, R.
Endocrinology 132, 2254-2261, 1993
A;Title: Purification and characterization of a 3,5,3'-L-triiodothyronine-specific bi
A;Reference number: A49192; MUID:93238653
A;Accession: A49192
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <YAM>
A;Experimental source: tadpole plasma
A;Note: sequence extracted from NCBI backbone (NCBIP:130235)


  Query Match             24.0%;  Score 23;  DB 2;  Length 19;
  Best Local Similarity   45.5%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLSLLD 18
          ||  | : :||
Db      7 SKXPLMVKVLD 17
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:06 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-11
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	31.2	18	2	S14560	probable heme-bind
2	28	29.2	14	2	S13864	methyl coenzyme M
3	26	27.1	14	4	I52618	hemoglobin beta ch
4	25	26.0	6	2	A61411	ameletin - rat
5	24	25.0	17	2	I49593	cystic fibrosis tr
6	24	25.0	17	2	I84733	gene CFTR protein
7	23	24.0	19	2	A49192	transthyretin - bu
8	23	24.0	20	2	S39049	cytotoxin-binding
9	22	22.9	13	2	PH0786	T-cell receptor al
10	22	22.9	15	2	S65717	prostaglandin D-sy
11	22	22.9	15	2	PA0046	protein QA100044 -
12	22	22.9	16	2	S34444	blaZ protein - Sta
13	22	22.9	18	2	S54272	CTC 75 protein - h
14	21	21.9	13	2	G22565	R-phycoerythrin ga
15	21	21.9	14	2	S29209	avenin alpha-2 - o
16	21	21.9	14	2	A44920	2-halobenzoate 1,2
17	21	21.9	15	2	PN0662	dystrophin-associa
18	21	21.9	16	2	B45895	T-cell surface gly
19	21	21.9	17	2	S32587	L-ascorbate peroxi
20	21	21.9	17	2	A35550	adrenocortical cel
21	21	21.9	19	2	S59717	hypothetical prote
22	21	21.9	20	2	H49034	nuclear antigen EB
23	20	20.8	15	2	S03955	acidic fibroblast
24	20	20.8	16	2	S01669	rRNA N-glycosidase
25	20	20.8	16	2	S02473	coat protein VP1 -
26	20	20.8	16	2	S70331	endosperm protein,
27	20	20.8	18	2	S45373	translation elonga
28	20	20.8	20	2	S29108	glutathione transf
29	20	20.8	20	2	S63602	glutathione S-tran
30	20	20.8	20	2	PL0192	Ig lambda 2 chain
31	20	20.8	20	2	A44921	hydroxypyruvate re
32	20	20.8	20	2	S46479	retinoid-X-recepto
33	19	19.8	9	3	PC7076	spectrin alpha cha
34	19	19.8	10	2	A13687	caerulein-like pep
35	19	19.8	12	2	S26556	T-cell receptor be
36	19	19.8	13	2	S47365	T-cell antigen rec
37	19	19.8	14	2	S14336	mastoparan B - hor
38	19	19.8	14	2	A47421	leukotriene B-4 12
39	19	19.8	15	2	S36890	ribosomal protein
40	19	19.8	17	2	S61451	hypothetical prote
41	19	19.8	17	2	PH1331	Ig heavy chain DJ
42	19	19.8	18	2	B35910	neurofibromatosis-
43	19	19.8	18	2	S48863	cyclin C - mouse (
44	19	19.8	19	2	S43652	serpin WSZCI - whe
45	19	19.8	19	2	A47689	flagellar sheath p

ALIGNMENTS

RESULT 1
S14560
probable heme-binding protein - garden pea chloroplast (fragment)
C;Species: chloroplast Pisum sativum (garden pea)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C;Accession: S14560
R;Smith, A.G.; Wilson, R.J.; Kaethner, T.M.; Willey, D.L.; Gray, J.C.
submitted to the EMBL Data Library, October 1990
A;Reference number: S14557
A;Accession: S14560
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <SMI>
A;Cross-references: EMBL:X54750; NID:g12194; PID:g12198
C;Genetics:
A;Genome: chloroplast
C;Superfamily: maize chloroplast protein cemA
C;Keywords: chloroplast; heme; transmembrane protein

Query Match 31.2%; Score 30; DB 2; Length 18;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 SKKKLFLSLL 17
:||| |: ||
Db 2 AKKKAFIPLL 11

RESULT 2
S13864
methyl coenzyme M reductase (EC 1.8.-.-) II alpha chain - Methanobacterium thermoauto
C;Species: Methanobacterium thermoautotrophicum
A;Variety: strain Marburg
C;Date: 19-Mar-1997 #sequence_revision 13-Sep-1998 #text_change 30-Oct-1998
C;Accession: S13864
R;Rospert, S.; Linder, D.; Ellermann, J.; Thauer, R.K.
Eur. J. Biochem. 194, 871-877, 1990
A;Title: Two genetically distinct methyl-coenzyme M reductases in Methanobacterium th
A;Reference number: S13864; MUID:91099370
A;Accession: S13864
A;Molecule type: protein
A;Residues: 1-14 <ROS>
A;Experimental source: strain Marburg
C;Keywords: methanogenesis; oxidoreductase

Query Match 29.2%; Score 28; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 KKKLFLSLL 17

```
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-305-871A-9


  Query Match             27.1%;  Score 26;  DB 1;  Length 16;
  Best Local Similarity   55.6%;  Pred. No. 2.3e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
             ::| | |||
Db       8 LKAIYEKKK 16


Search completed: December 21, 2000, 08:31:49
Job time: 372 sec
```

```
;    APPLICANT:  Kourilsky, Phillipe
;    TITLE OF INVENTION:  Altered Major Histocompatibility Complex
;    TITLE OF INVENTION:
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,985B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  03495.0106-04000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-481-985B-26


  Query Match             27.1%;  Score 26;  DB 3;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           ::| | |||
Db       7 LKAIYEKKK 15


RESULT  14
PCT-US95-01671-2
; Sequence 2, Application PC/TUS9501671
;  GENERAL INFORMATION:
;    APPLICANT:  University of Rochester
;    TITLE OF INVENTION:  Inducing Tolerance With Tolerogenic
;    TITLE OF INVENTION:  Fusion Proteins
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schwegman, Lundberg & Woessner
;      STREET:  3500 IDS Center
;      CITY:  Minneapolis
;      STATE:  Minnesota
;      COUNTRY:  United States of America
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/01671
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kevin W. Raasch
;      REGISTRATION NUMBER:  35,651
;      REFERENCE/DOCKET NUMBER:  850.114WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-339-0331
;      TELEFAX:  612-339-3061
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-01671-2


  Query Match             27.1%;  Score 26;  DB 4;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           ::| | |||
Db       7 LKAIYEKKK 15


RESULT  15
US-08-305-871A-9
; Sequence 9, Application US/08305871A
; Patent No. 5736142
;  GENERAL INFORMATION:
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Gaeta, Federico
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sidney, John
;    APPLICANT:  Alexander, Jeffrey L.
;    TITLE OF INVENTION:  Alteration of Immune Response Using Pan
;    TITLE OF INVENTION:  DR-Binding Peptides
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/305,871A
;      FILING DATE:  14-SEP-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/121,101
;      FILING DATE:  14-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  14137-0062-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  IRI44PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-06077-5


  Query Match             27.1%;  Score 26;  DB 4;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 SYSKKK 11
             || |||
Db       1 SYGKKK 6


RESULT  11
US-08-195-874-2
; Sequence 2, Application US/08195874
; Patent No. 5817308
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, David W.
;    APPLICANT:  Zambidis, Elias T.
;    TITLE OF INVENTION:  Tolerogenic Fusion Proteins of
;    TITLE OF INVENTION:  Immunoglobulins and Methods for Inducing and Maintaining
;    TITLE OF INVENTION:  Tolerance
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schwegman, Lundberg & Woessner, P.A.
;      STREET:  3500 IDS Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/195,874
;      FILING DATE:  02-FEB-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mueting, Ann M.
;      REGISTRATION NUMBER:  33,977
;      REFERENCE/DOCKET NUMBER:  850.14-US-01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-339-0331
;      TELEFAX:  612-339-3061
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-195-874-2


  Query Match             27.1%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
             ::| | |||
Db       7 LKAIYEKKK 15


RESULT  12
US-08-484-905-26
; Sequence 26, Application US/08484905
; Patent No. 5976551
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
;    APPLICANT:  Kourilsky, Philippe
;    TITLE OF INVENTION:  An Altered Major Histocompatibility
;    TITLE OF INVENTION:  Complex(MHC) Determinant and Methods for Using the
;    TITLE OF INVENTION:  Determinant
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS-/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,905
;      FILING DATE:  07-JUNE-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Potter, Jane E. R.
;      REGISTRATION NUMBER:  33,332
;      REFERENCE/DOCKET NUMBER:  03495.0106-03000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-905-26


  Query Match             27.1%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
             ::| | |||
Db       7 LKAIYEKKK 15


RESULT  13
US-08-481-985B-26
; Sequence 26, Application US/08481985B
; Patent No. 6011146
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
```

```
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLSL 16
          ||||| | :
Db      1 SKKKLALCI 9


RESULT   8
US-08-454-207A-41
; Sequence 41, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-41


  Query Match             27.1%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLS 15
          ||||| |:
Db      1 SKKKLALT 8


RESULT   9
US-08-454-207A-44
; Sequence 44, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-44


  Query Match             27.1%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   85.7%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 SKKKLFL 14
          ||||| |
Db      1 SKKKLHL 7


RESULT  10
PCT-US95-06077-5
; Sequence 5, Application PC/TUS9506077
;  GENERAL INFORMATION:
;    APPLICANT:  Immunobiology Research, Institute Inc.
;    TITLE OF INVENTION:  Vaccine Interdiction of Extracellular
;    TITLE OF INVENTION:  Transactivating Proteins of Human Immunodeficiency Virus
;    TITLE OF INVENTION:  and Other Chronically Infecting Viruses Employing Similar
;    TITLE OF INVENTION:  Intercellular Transactivating Strategies
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06077
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/247,991
;      FILING DATE:  23-MAY-1994
```

; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/654,839
; FILING DATE: 19910213
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/252,753
; FILING DATE: 03-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/415,029
; FILING DATE: 29-SEP-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Bingham, Douglas A
; REGISTRATION NUMBER: 32,457
; REFERENCE/DOCKET NUMBER: SCR0367P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-07-654-839-5

Query Match 29.2%; Score 28; DB 1; Length 19;
Best Local Similarity 38.5%; Pred. No. 1.3e+02;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 7 YSKKKLFLSLLDF 19
|| || : :: |
Db 1 YSMKKTTMKIIPF 13

RESULT 6
US-08-454-207A-45
; Sequence 45, Application US/08454207A
; Patent No. 5710123
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF SELECTIN BINDING
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,207A
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12110
; FILING DATE: 13-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/997,771
; FILING DATE: 18-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: CCOR-0183
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-454-207A-45

Query Match 28.1%; Score 27; DB 1; Length 9;
Best Local Similarity 85.7%; Pred. No. 1.2e+05;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 SKKKLFL 14
||||| |
Db 1 SKKKLIL 7

RESULT 7
US-08-454-207A-30
; Sequence 30, Application US/08454207A
; Patent No. 5710123
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF SELECTIN BINDING
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,207A
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12110
; FILING DATE: 13-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/997,771
; FILING DATE: 18-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: CCOR-0183
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-454-207A-30

Query Match 27.1%; Score 26; DB 1; Length 9;
Best Local Similarity 66.7%; Pred. No. 1.2e+05;

```
          |||||||
Db      1 SKKKLFL 7


RESULT   2
5266328-11
;Patent No. 5266328
;     APPLICANT: SKUBITZ, AMY P.N.;FURCHT, LEO T.
;     TITLE OF INVENTION: LAMININ  CHAIN POLYPEPTIDES FROM
;THE CARBOXY TERMINAL GLOBULAR DOMAIN
;     NUMBER OF SEQUENCES: 14
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/573,672
;       FILING DATE: 27-AUG-1990
;SEQ ID NO:11:
;       LENGTH: 17
5266328-11


  Query Match             33.3%;  Score 32;  DB 5;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TVQASYSKKKLF 13
          ||:  | |:| |
Db      6 TVKTEYIKRKAF 17


RESULT   3
US-09-100-414B-26
; Sequence 26, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;     APPLICANT:  Wang, Chang Yi
;     TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;     TITLE OF INVENTION:  IMMUNOGENS
;     NUMBER OF SEQUENCES:  106
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Morgan & Finnegan, L.L.P.
;       STREET:  345 Park Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10154-0054
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC Windows
;       SOFTWARE:  Word 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/100,414B
;       FILING DATE:  20-JUNE-1998
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4157
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-758-4800
;       TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  26:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-26


  Query Match             30.2%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 KKKLFL 14
          ||||||
Db      1 KKKLFL 6


RESULT   4
US-09-100-414B-30
; Sequence 30, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;     APPLICANT:  Wang, Chang Yi
;     TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;     TITLE OF INVENTION:  IMMUNOGENS
;     NUMBER OF SEQUENCES:  106
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Morgan & Finnegan, L.L.P.
;       STREET:  345 Park Avenue
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10154-0054
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC Windows
;       SOFTWARE:  Word 97
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/100,414B
;       FILING DATE:  20-JUNE-1998
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Maria H. Lin
;       REGISTRATION NUMBER:  29,323
;       REFERENCE/DOCKET NUMBER:  1151-4157
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  212-758-4800
;       TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  30:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  18 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-09-100-414B-30


  Query Match             30.2%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 KKKLFL 14
          ||||||
Db      1 KKKLFL 6


RESULT   5
US-07-654-839-5
; Sequence 5, Application US/07654839
; Patent No. 5372933
;  GENERAL INFORMATION:
;     APPLICANT:  Zamarron, Concepcion
;     APPLICANT:  Plow, Edward F
;     APPLICANT:  Ginsberg, Mark H
;     TITLE OF INVENTION:  MONOCLONAL ANTIBODIES AGAINST
;     TITLE OF INVENTION:  RECEPTOR-INDUCED BINDING SITES
;     NUMBER OF SEQUENCES:  9
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  The Scripps Research Institute, Office of
;       ADDRESSEE:  Patent Counsel
;       STREET:  10666 No. 5372933th Torrey Pines Road., Suite 220,
;       STREET:  Mail Drop TPC8
;       CITY:  La Jolla
;       STATE:  CA
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:49 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-11
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	34.4	9	1	US-08-454-207A-43	Sequence 43, Appl
2	32	33.3	17	5	5266328-11	Patent No. 5266328
3	29	30.2	18	3	US-09-100-414B-26	Sequence 26, Appl
4	29	30.2	18	3	US-09-100-414B-30	Sequence 30, Appl
5	28	29.2	19	1	US-07-654-839-5	Sequence 5, Appli
6	27	28.1	9	1	US-08-454-207A-45	Sequence 45, Appl
7	26	27.1	9	1	US-08-454-207A-30	Sequence 30, Appl
8	26	27.1	9	1	US-08-454-207A-41	Sequence 41, Appl
9	26	27.1	9	1	US-08-454-207A-44	Sequence 44, Appl
10	26	27.1	14	4	PCT-US95-06077-5	Sequence 5, Appli
11	26	27.1	15	2	US-08-195-874-2	Sequence 2, Appli
12	26	27.1	15	2	US-08-484-905-26	Sequence 26, Appl
13	26	27.1	15	3	US-08-481-985B-26	Sequence 26, Appl
14	26	27.1	15	4	PCT-US95-01671-2	Sequence 2, Appli
15	26	27.1	16	1	US-08-305-871A-9	Sequence 9, Appli
16	26	27.1	20	1	US-08-218-025A-81	Sequence 81, Appl
17	25	26.0	9	1	US-08-454-207A-29	Sequence 29, Appl
18	25	26.0	9	1	US-08-454-207A-31	Sequence 31, Appl
19	25	26.0	9	1	US-08-454-207A-32	Sequence 32, Appl
20	25	26.0	9	1	US-08-454-207A-33	Sequence 33, Appl
21	25	26.0	9	1	US-08-454-207A-34	Sequence 34, Appl
22	25	26.0	9	1	US-08-454-207A-35	Sequence 35, Appl
23	25	26.0	9	1	US-08-454-207A-36	Sequence 36, Appl
24	25	26.0	9	1	US-08-454-207A-37	Sequence 37, Appl
25	25	26.0	9	1	US-08-454-207A-38	Sequence 38, Appl
26	25	26.0	9	1	US-08-454-207A-39	Sequence 39, Appl
27	25	26.0	9	1	US-08-454-207A-40	Sequence 40, Appl
28	25	26.0	9	3	US-08-159-339A-78	Sequence 78, Appl
29	25	26.0	10	1	US-08-454-207A-2	Sequence 2, Appli
30	25	26.0	10	1	US-08-454-207A-5	Sequence 5, Appli
31	25	26.0	10	2	US-08-456-112B-42	Sequence 42, Appl
32	25	26.0	10	3	US-08-159-339A-77	Sequence 77, Appl
33	25	26.0	11	2	US-08-310-912A-117	Sequence 117, App
34	25	26.0	11	4	PCT-US95-04589-117	Sequence 117, App
35	25	26.0	13	4	PCT-US94-01234-47	Sequence 47, Appl
36	25	26.0	14	2	US-09-133-774-9	Sequence 9, Appli
37	25	26.0	14	3	US-09-303-862-9	Sequence 9, Appli
38	25	26.0	17	1	US-07-992-288-4	Sequence 4, Appli
39	25	26.0	17	1	US-07-989-764-4	Sequence 4, Appli
40	25	26.0	18	1	US-08-233-203-10	Sequence 10, Appl
41	25	26.0	18	1	US-08-182-483A-25	Sequence 25, Appl
42	25	26.0	18	1	US-08-243-879A-24	Sequence 24, Appl
43	25	26.0	18	1	US-08-499-523-45	Sequence 45, Appl
44	25	26.0	18	3	US-09-100-414B-24	Sequence 24, Appl
45	25	26.0	20	2	US-08-564-972-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-454-207A-43
; Sequence 43, Application US/08454207A
; Patent No. 5710123
; GENERAL INFORMATION:
; APPLICANT: Heavner, George A.
; APPLICANT: Kruszynski, Marian
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF SELECTIN BINDING
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,207A
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/12110
; FILING DATE: 13-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/997,771
; FILING DATE: 18-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: CCOR-0183
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-454-207A-43

Query Match 34.4%; Score 33; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SKKKLFL 14

```
;       APPLICATION NUMBER:  US/08/596,257A
;       FILING DATE:  08-FEB-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  DE 4327165.0
;       FILING DATE:  09-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meilman, Edward A.
;       REGISTRATION NUMBER:  24,735
;       REFERENCE/DOCKET NUMBER:  P/951-118
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 382-0700
;       TELEFAX:  (212) 382-0888
;       TELEX:  236925
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  solanum tuberosum
US-08-596-257A-10


  Query Match             24.3%;  Score 27;  DB 3;  Length 15;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPEHLL 14
           | | ::| |: ::
Db       1 LEQTNYGLPQQVI 13


RESULT  15
US-08-860-339-10
; Sequence 10, Application US/08860339
; Patent No. 6117665
; GENERAL INFORMATION:
;  APPLICANT: Kossmann, Jens
;  APPLICANT: Emmermann, Michael
;  APPLICANT: Virgin, Ivar
;  APPLICANT: Renz, Andreas
;  TITLE OF INVENTION: DNA MOLECULES CODING FOR DEBRANCHING ENZYMES DERIVED
;  TITLE OF INVENTION: FROM PLANTS
;  FILE REFERENCE: AGREVO-6
;  CURRENT APPLICATION NUMBER: US/08/860,339
;  CURRENT FILING DATE: 1997-11-25
;  EARLIER APPLICATION NUMBER: DE P4447387.7
;  EARLIER FILING DATE: 1994-12-22
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Solanum tuberosum
US-08-860-339-10


  Query Match             24.3%;  Score 27;  DB 3;  Length 15;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPEHLL 14
           | | ::| |: ::
Db       1 LEQTNYGLPQQVI 13


Search completed: December 21, 2000, 08:36:43
Job time: 387 sec
```

```
Db      7 ERLLEDLLKALN 18


RESULT  12
US-08-940-096-134
; Sequence 134, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,096
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0005-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | |::|:
Db       7 ERLLEDLLKALN 18


RESULT  13
US-08-765-783A-83
; Sequence 83, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-83


  Query Match             24.3%;  Score 27;  DB 2;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 QMDFGFP 10
            |  ||||
Db       1 QHHFGFP 7


RESULT  14
US-08-596-257A-10
; Sequence 10, Application US/08596257A
; Patent No. 6001628
;  GENERAL INFORMATION:
;    APPLICANT:  KOSSMAN, Jens
;    APPLICANT:  EMMERMANN, Michael
;    APPLICANT:  VIRGIN, Ivar
;    TITLE OF INVENTION:  DEBRANCHING ENZYMES AND DNA SEQUENCES
;    TITLE OF INVENTION:  CODING THEM, SUITABLE FOR CHANGING THE DEGREE OF BRANCHING
;    TITLE OF INVENTION:  OF AMYLOPECTIN STARCH IN PLANTS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ostrolenk, Faber, Gerb & Soffen
;      STREET:  1180 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10036-8403
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  401:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-401


  Query Match             25.2%;  Score 28;  DB 3;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 61;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 FPEHLL 14
             :||||:
Db       2 YPEHLV 7


RESULT  10
US-08-940-095-134
; Sequence 134, Application US/08940095
; Patent No. 6004925
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    APPLICANT:  Dufourcq, Jean
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,095
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0004-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6004925e
US-08-940-095-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
             | || | |::|:
Db       7 ERLLEDLLKALN 18


RESULT  11
US-08-940-093-134
; Sequence 134, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,093
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0006-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6037323e
US-08-940-093-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
             | || | |::|:
```

```
;       ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;       STREET:  Four Embarcadero Center, Suite 3400
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111-4187
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/827,618A
;       FILING DATE:  09-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,725
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/716,909
;       FILING DATE:  18-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/586,536
;       FILING DATE:  21-SEP-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Trecartin, Richard F.
;       REGISTRATION NUMBER:  31,801
;       REFERENCE/DOCKET NUMBER:  A-60780-12/RFT/MTK
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-827-618A-22


  Query Match             26.1%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 93;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   8
US-08-483-952A-22
; Sequence 22, Application US/08483952A
; Patent No. 6011139
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,952A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,859
;      FILING DATE:  17-SEP-1993
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/716,909
;      FILING DATE:  18-JUN-1991
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/586,536
;      FILING DATE:  21-SEP-1990
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-7/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-952A-22


  Query Match             26.1%;  Score 29;  DB 3;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 93;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   9
US-08-159-339A-401
; Sequence 401, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
```

```
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  System 7.0.1
;       SOFTWARE:  Microsoft Word 5.0a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/199,508
;       FILING DATE:  February 18, 1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/021,536
;       FILING DATE:  February 23, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5717058ak, Henry P.
;       REGISTRATION NUMBER:  33200
;       REFERENCE/DOCKET NUMBER:  121 CIP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  303-541-3322
;       TELEFAX:  303-444-3013
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20
;       TYPE:  amino acid
;       TOPOLOGY:  unknown to applicant
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  yes
US-08-199-508-2


  Query Match             27.0%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 63;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFLQ 19
           ||:| ::  | :|
Db       1 FGYPVYVFGDCVQ 13


RESULT   5
US-08-399-696-85
; Sequence 85, Application US/08399696
; Patent No. 5756669
;  GENERAL INFORMATION:
;     APPLICANT:
;     TITLE OF INVENTION:  p53-BINDING POLYPEPTIDES AND
;     TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING SAME
;     NUMBER OF SEQUENCES:  126
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/399,696
;       FILING DATE:  02-MAR-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/156,671
;       FILING DATE:  22-NOV-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Smith, William M
;       REGISTRATION NUMBER:  30,223
;       REFERENCE/DOCKET NUMBER:  15522-000710
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 326-2400
;       TELEFAX:  (415) 326-2422
;   INFORMATION FOR SEQ ID NO:  85:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  19 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  unknown
;       TOPOLOGY:  unknown
;     MOLECULE TYPE:  peptide
US-08-399-696-85


  Query Match             26.1%;  Score 29;  DB 1;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 87;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFL 18
           |||::| |
Db      12 EHLIIDGL 19


RESULT   6
US-08-484-530-22
; Sequence 22, Application US/08484530
; Patent No. 5846740
;  GENERAL INFORMATION:
;     APPLICANT:  Tobin, Allan J
;     APPLICANT:  Erlander, Mark G
;     APPLICANT:  Kaufman, Daniel L.
;     TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;     NUMBER OF SEQUENCES:  60
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;       STREET:  Four Embarcadero Center, Suite 3400
;       CITY:  San Francisoc
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111-4187
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/484,530
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Trecartin, Richard F.
;       REGISTRATION NUMBER:  31,801
;       REFERENCE/DOCKET NUMBER:  A-60780-8/RFT/MTK
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-484-530-22


  Query Match             26.1%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 93;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   7
US-08-827-618A-22
; Sequence 22, Application US/08827618A
; Patent No. 5998366
;  GENERAL INFORMATION:
;     APPLICANT:  Tobin, Allan J
;     APPLICANT:  Erlander, Mark G
;     APPLICANT:  Kaufman, Daniel L.
;     TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;     NUMBER OF SEQUENCES:  60
;     CORRESPONDENCE ADDRESS:
```

```
  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   2
US-08-940-093-133
; Sequence 133, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,093
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0006-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6037323e
US-08-940-093-133


  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   3
US-08-940-096-133
; Sequence 133, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,096
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0005-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-133


  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   4
US-08-199-508-2
; Sequence 2, Application US/08199508
; Patent No. 5717058
;  GENERAL INFORMATION:
;    APPLICANT:  Matthews, Maura-Ann H.
;    APPLICANT:  Stetler, Gary L.
;    APPLICANT:  Anthony-Cahill, Spencer J.
;    APPLICANT:  Anderson, David C.
;    TITLE OF INVENTION:  Modulators of Gene Expression
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  5797 Central Avenue
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:43 ; Search time 25.87 Seconds
(without alignments)
14.254 Million cell updates/sec

Title: US-08-934-367-10
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 92835

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	28.8	22	3	US-08-940-095-133	Sequence 133, App
2	32	28.8	22	3	US-08-940-093-133	Sequence 133, App
3	32	28.8	22	3	US-08-940-096-133	Sequence 133, App
4	30	27.0	20	1	US-08-199-508-2	Sequence 2, Appli
5	29	26.1	19	1	US-08-399-696-85	Sequence 85, Appl
6	29	26.1	20	2	US-08-484-530-22	Sequence 22, Appl
7	29	26.1	20	2	US-08-827-618A-22	Sequence 22, Appl
8	29	26.1	20	3	US-08-483-952A-22	Sequence 22, Appl
9	28	25.2	10	3	US-08-159-339A-401	Sequence 401, App
10	28	25.2	22	3	US-08-940-095-134	Sequence 134, App
11	28	25.2	22	3	US-08-940-093-134	Sequence 134, App
12	28	25.2	22	3	US-08-940-096-134	Sequence 134, App
13	27	24.3	9	2	US-08-765-783A-83	Sequence 83, Appl
14	27	24.3	15	3	US-08-596-257A-10	Sequence 10, Appl
15	27	24.3	15	3	US-08-860-339-10	Sequence 10, Appl
16	27	24.3	20	1	US-07-678-974D-8	Sequence 8, Appli
17	27	24.3	20	2	US-08-162-149-10	Sequence 10, Appl
18	27	24.3	20	2	US-08-945-168-13	Sequence 13, Appl
19	27	24.3	22	2	US-08-559-524A-12	Sequence 12, Appl
20	27	24.3	22	3	US-08-749-707-12	Sequence 12, Appl
21	26	23.4	7	1	US-07-923-724-47	Sequence 47, Appl
22	26	23.4	7	2	US-08-609-426A-47	Sequence 47, Appl
23	26	23.4	7	2	US-08-374-652C-37	Sequence 37, Appl
24	26	23.4	11	1	US-08-039-778B-6	Sequence 6, Appli
25	26	23.4	12	1	US-08-555-394-13	Sequence 13, Appl
26	26	23.4	12	3	US-08-745-892-13	Sequence 13, Appl
27	26	23.4	13	1	US-08-620-213-3	Sequence 3, Appli
28	26	23.4	16	1	US-08-346-455B-52	Sequence 52, Appl
29	26	23.4	16	3	US-08-977-221-52	Sequence 52, Appl
30	26	23.4	16	4	PCT-US95-06613-52	Sequence 52, Appl
31	26	23.4	17	1	US-08-620-213-1	Sequence 1, Appli
32	26	23.4	19	1	US-08-238-163-20	Sequence 20, Appl
33	26	23.4	19	1	US-07-977-696C-36	Sequence 36, Appl
34	26	23.4	19	1	US-08-129-930B-36	Sequence 36, Appl
35	26	23.4	22	1	US-08-039-778B-5	Sequence 5, Appli
36	25	22.5	6	1	US-08-487-006-76	Sequence 76, Appl
37	25	22.5	6	2	US-08-488-659A-76	Sequence 76, Appl
38	25	22.5	8	3	US-08-582-776C-29	Sequence 29, Appl
39	25	22.5	8	3	US-08-434-831B-29	Sequence 29, Appl
40	25	22.5	10	2	US-08-318-157B-25	Sequence 25, Appl
41	25	22.5	11	1	US-08-039-778B-3	Sequence 3, Appli
42	25	22.5	13	3	US-09-140-149-21	Sequence 21, Appl
43	25	22.5	15	1	US-07-918-181A-17	Sequence 17, Appl
44	25	22.5	15	1	US-07-918-181A-21	Sequence 21, Appl
45	25	22.5	15	1	US-08-080-073-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-08-940-095-133
; Sequence 133, Application US/08940095
; Patent No. 6004925
; GENERAL INFORMATION:
;   APPLICANT: Dasseux, Jean-Louis
;   APPLICANT: Sekul, Renate
;   APPLICANT: Buttner, Klaus
;   APPLICANT: Cornut, Isabelle
;   APPLICANT: Metz, Gunther
;   APPLICANT: Dufourcq, Jean
;   TITLE OF INVENTION: APOLIPOPROTEIN A-I AGONISTS
;   TITLE OF INVENTION: AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;   NUMBER OF SEQUENCES: 258
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Pennie & Edmonds LLP
;     STREET: 1155 Avenue of the Americas
;     CITY: New York
;     STATE: NY
;     COUNTRY: USA
;     ZIP: 10036-2811
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/940,095
;     FILING DATE: 29-SEP-1997
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Coruzzi, Laura A
;     REGISTRATION NUMBER: 30,742
;     REFERENCE/DOCKET NUMBER: 009196-0004-999
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 650-493-4935
;     TELEFAX: 650-493-5556
;     TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 133:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 22 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: No. 6004925e
US-08-940-095-133
```

```
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 10-Dec-1993 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000
C;Accession: S36662
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of th
A;Reference number: S21152; MUID:92339502
A;Accession: S36662
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>


  Query Match             20.7%;  Score 23;  DB 2;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFPE 11
           ||:|:
Db       3 FGYPK 7


RESULT  15
PA0015
seed storage protein 12S 2 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 11-Jul-1997
C;Accession: PA0015
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensiona
A;Reference number: PA0001
A;Accession: PA0015
A;Molecule type: protein
A;Residues: 1-14 <KAM>
A;Experimental source: seed
C;Keywords: pyroglutamic acid; seed; storage protein
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental


  Query Match             20.7%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 1e+03;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 QMDFGFPEHL 13
           |   | || |
Db       1 QQYLGVPEQL 10



Search completed: December 21, 2000, 08:37:13
Job time: 317 sec
```

```
  Best Local Similarity   83.3%;  Pred. No. 6.9e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PEHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT   9
A60834
angiotensin I precursor - dog (fragment)
N;Alternate names: angiotensinogen I
N;Contains: angiotensin I
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Mar-1998
C;Accession: A60834
R;Oliver, J.A.
Hypertension 11, 21-27, 1988
A;Title: Purification and partial characterization of canine angiotensinogen.
A;Reference number: A60834; MUID:88113996
A;Accession: A60834
A;Molecule type: protein
A;Residues: 1-15 <OLI>
C;Superfamily: antithrombin III
C;Keywords: glycoprotein; plasma
F;1-10/Product: angiotensin I #status predicted <MAT>


  Query Match             21.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PEHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT  10
B31769
T-cell receptor delta-2 chain J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
C;Accession: B31769
R;Loh, E.Y.; Cwirla, S.; Serafini, A.T.; Phillips, J.H.; Lanier, L.L.
Proc. Natl. Acad. Sci. U.S.A. 85, 9714-9718, 1988
A;Title: Human T-cell-receptor delta chain: genomic organization, diversity, and express
A;Reference number: A94221; MUID:89071766
A;Accession: B31769
A;Molecule type: DNA
A;Residues: 1-17 <LOH>
A;Cross-references: GB:L36386; NID:g540455; PIDN:AAA61108.1; PID:g540456
C;Keywords: T-cell receptor


  Query Match             21.6%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   37.5%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPEHLLVD 16
           |  |: ||    |:|:
Db       1 LTAQLFFGKGTQLIVE 16


RESULT  11
PN0175
glutathione transferase (EC 2.5.1.18) 1 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 22-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 15-Jun-1996
C;Accession: PN0175
R;Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.
submitted to JIPID, December 1995
A;Description: Two dimensional electrophoresis of plant proteins and standardization of
A;Reference number: PN0173
A;Accession: PN0175
A;Molecule type: protein
A;Residues: 1-18 <TSU>
A;Experimental source: leaf
C;Keywords: transferase


  Query Match             21.6%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   80.0%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 DFGFP 10
           ||| |
Db       5 DFGHP 9


RESULT  12
H75063
hypothetical protein PAB7382 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: H75063
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A;Reference number: A75001
A;Accession: H75063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <KAW>
A;Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50405.1; PID:e151
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB7382


  Query Match             21.6%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   35.7%;  Pred. No. 9.2e+02;
  Matches    5;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPEHLLVD 16
           |::: |  : :|||
Db       3 LRVETGCIDRVLVD 16


RESULT  13
A37984
ADP,ATP carrier protein - yeast (Candida parapsilosis) (fragment)
C;Species: Candida parapsilosis
C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 07-Jun-1996
C;Accession: A37984
R;Guerin, B.; Bukusoglu, C.; Rakotomanana, F.; Wohlrab, H.
J. Biol. Chem. 265, 19736-19741, 1990
A;Title: Mitochondrial phosphate transport. N-ethylmaleimide insensitivity correlates
A;Reference number: A37984; MUID:91060585
A;Accession: A37984
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <GUE>


  Query Match             21.6%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 HLLVDFL 18
           : |:|||
Db       5 NFLIDFL 11


RESULT  14
S36662
dermorphin (Lys-7) [validated] - two-colored leaf frog
```

S59492
formate dehydrogenase alpha chain - Alcaligenes eutrophus (fragment)
C;Species: Alcaligenes eutrophus
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59492
R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.
Biol. Chem. Hoppe-Seyler 376, 561-568, 1995
A;Title: Structural and immunological studies on the soluble formate dehydrogenase from
A;Reference number: S59492; MUID:96145736
A;Accession: S59492
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRI>

Query Match 23.4%; Score 26; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 3.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QMDFGFP 10
::||| |
Db 6 EIDFGTP 12

RESULT 4
S71306
heat shock protein 90 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S71306
R;Conconi, M.; Szweda, L.I.; Levine, R.L.; Stadtman, E.R.; Friguet, B.
Arch. Biochem. Biophys. 331, 232-240, 1996
A;Title: Age-related decline of rat liver multicatalytic proteinase activity and protect
A;Reference number: S71306; MUID:96299287
A;Accession: S71306
A;Molecule type: protein
A;Residues: 1-15 <CON>
A;Experimental source: liver
C;Keywords: heat shock; phosphoprotein; stress-induced protein

Query Match 22.5%; Score 25; DB 2; Length 15;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 10 PEHLLVDFLQ 19
|:| :|: |:
Db 6 PDHPIVETLR 15

RESULT 5
T44936
calmodulin kinase 2 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44936
R;Alemany, V.; Aligue, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z22873
A;Accession: T44936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-16 <ALE>
A;Cross-references: EMBL:U57982; PIDN:AAD09466.1

Query Match 22.5%; Score 25; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FGFP 10
||||
Db 5 FGFP 8

RESULT 6
B61597
cytochrome P450 AL-2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: B61597
R;Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.
Drug Metab. Dispos. 19, 291-297, 1991
A;Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto
A;Reference number: A61597; MUID:91292910
A;Accession: B61597
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <SHI>

Query Match 22.1%; Score 24.5; DB 2; Length 14;
Best Local Similarity 47.1%; Pred. No. 5.7e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 LLLQMDFGFPEHLLVDF 17
||| : | ||| |
Db 3 LLLSLSF-----LLVGF 14

RESULT 7
S03879
6-phosphofructokinase (EC 2.7.1.11) B - rabbit (fragment)
N;Alternate names: phosphofructo-1-kinase B
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 28-Apr-1993
C;Accession: S03879
R;Valaitis, A.P.; Foe, L.G.; Kwiatkowska, D.; Latshaw, S.P.; Kemp, R.G.
Biochim. Biophys. Acta 995, 187-194, 1989
A;Title: The sites of phosphorylation of rabbit brain phosphofructo-1-kinase by cycli
A;Reference number: S03878; MUID:89194250
A;Accession: S03879
A;Molecule type: protein
A;Residues: 1-13 <VAL>
A;Note: the sequence from the summary is inconsistent with that from table I and Fig.
C;Keywords: glycolysis; phosphotransferase

Query Match 21.6%; Score 24; DB 2; Length 13;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LQMDFGF 9
| || ||
Db 7 LSMDKGF 13

RESULT 8
A01250
angiotensin precursor - horse (fragment)
C;Species: Equus caballus (domestic horse)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 20-Mar-1998
C;Accession: A92775; A01250
R;Skeggs Jr., L.T.; Kahn, J.R.; Lentz, K.; Shumway, N.P.
J. Exp. Med. 106, 439-453, 1957
A;Reference number: A92775
A;Accession: A92775
A;Molecule type: protein
A;Residues: 1-14 <SKE>
C;Superfamily: antithrombin III
C;Keywords: blood pressure control; hormone; vasoconstrictor
F;1-10/Product: angiotensin I #status experimental <AN1>
F;1-8/Product: angiotensin II #status experimental <AN2>

Query Match 21.6%; Score 24; DB 2; Length 14;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:13 ; Search time 26.84 Seconds
(without alignments)
52.017 Million cell updates/sec

Title: US-08-934-367-10
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 4315

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.2	19	2	A39504	octamer-binding pr
2	27	24.3	21	2	A32521	hexokinase (EC 2.7
3	26	23.4	15	2	S59492	formate dehydrogen
4	25	22.5	15	2	S71306	heat shock protein
5	25	22.5	16	2	T44936	calmodulin kinase
6	24.5	22.1	14	2	B61597	cytochrome P450 AL
7	24	21.6	13	2	S03879	6-phosphofructokin
8	24	21.6	14	2	A01250	angiotensin precur
9	24	21.6	15	2	A60834	angiotensin I prec
10	24	21.6	17	2	B31769	T-cell receptor de
11	24	21.6	18	2	PN0175	glutathione transf
12	24	21.6	18	2	H75063	hypothetical prote
13	24	21.6	20	2	A37984	ADP,ATP carrier pr
14	23	20.7	7	2	S36662	dermorphin (Lys-7)
15	23	20.7	14	2	PA0015	seed storage prote
16	23	20.7	19	2	I46654	T-cell receptor de
17	23	20.7	21	2	T07683	proteinase inhibit
18	22	19.8	7	1	A61324	dermorphin - Rohde
19	22	19.8	11	2	PT0250	Ig heavy chain CRD
20	22	19.8	12	2	C36201	1-aminocyclopropan
21	22	19.8	14	2	A61002	photosystem II oxy
22	22	19.8	20	2	S72501	protein kinase C i
23	22	19.8	20	2	A31049	calsequestrin, fas
24	22	19.8	20	2	S77989	cytochrome-c oxida
25	22	19.8	20	2	A47105	dystroglycan - chi
26	22	19.8	20	2	S50203	zona pellucida gly
27	22	19.8	21	2	PX0078	alanine dehydrogen
28	21	18.9	9	2	B45796	dihydrolipoamide S
29	21	18.9	9	2	B20569	serum amyloid P-co
30	21	18.9	14	2	PA0104	protein QF200070 -
31	21	18.9	15	2	PH1788	T cell receptor al
32	21	18.9	15	2	S72432	epoxypropan isomer
33	21	18.9	17	2	A60743	ornithine carbamoy
34	21	18.9	18	2	I49408	cytochrome-c oxida
35	21	18.9	19	2	B60822	cytochrome P450 UT
36	21	18.9	19	2	D32071	T-cell receptor de
37	21	18.9	20	2	A60822	cytochrome P450 PB
38	21	18.9	20	2	S07232	ribulose-bisphosph
39	21	18.9	20	2	H49034	nuclear antigen EB
40	21	18.9	20	2	S35921	T-cell receptor ga
41	21	18.9	20	2	A54077	cytochrome b558 -
42	21	18.9	22	2	B54397	ubiquitin-carrier
43	21	18.9	22	2	D56978	collagen alpha 1(I
44	20	18.0	9	2	I46023	growth hormone rec
45	20	18.0	12	2	S56122	type I DNA methylt

ALIGNMENTS

RESULT 1
A39504
octamer-binding protein, Ku-like, 72K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39504
R;May, G.; Sutton, C.; Gould, H.
J. Biol. Chem. 266, 3052-3059, 1991
A;Title: Purification and characterization of Ku-2, an octamer-binding protein relate
A;Reference number: A39504; MUID:91131605
A;Accession: A39504
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <MAY>

Query Match 25.2%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 LQMDFGF 9
           |:|| ||
Db      11 LEMDVGF 17
```

RESULT 2
A32521
hexokinase (EC 2.7.1.1) I peptide III - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C;Accession: A32521
R;Schirch, D.M.; Wilson, J.E.
Arch. Biochem. Biophys. 257, 1-12, 1987
A;Title: Rat brain hexokinase: amino acid sequence at the substrate hexose binding si
A;Reference number: A90080; MUID:87324917
A;Accession: A32521
A;Molecule type: protein
A;Residues: 1-21 <SCH>
C;Superfamily: human hexokinase I; hexokinase homology
C;Keywords: ATP; glycolysis; phosphotransferase

Query Match 24.3%; Score 27; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       3 LQMDFGFPEH 12
           |   | || |
Db       3 LGFTFSFPXH 12
```

RESULT 3

Query Match 17.1%; Score 19; DB 1; Length 21;
Best Local Similarity 44.4%; Pred. No. 4.3e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       8 GFPEHLLVD 16
           | |   ||:
Db      12 GHPVEALVE 20
```

Search completed: December 21, 2000, 08:38:16
Job time: 149 sec

```
ID   UP21_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82027;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.1.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST
CC       L.MESENTERIODES, M.LUTEUS AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1927 MW;  328834D77BA353D2 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKKV 9


RESULT  13
UP25_UPEIN
ID   UP25_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82031;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.5.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST M.LUTEUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1940; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1941 MW;  5E94C6C757B463D9 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKGV 9


RESULT  14
SUCB_CANFA
ID   SUCB_CANFA     STANDARD;      PRT;    20 AA.
AC   P99507;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PUTATIVE SUCCINYL-COA LIGASE [GDP-FORMING] BETA-CHAIN (EC 6.2.1.4)
DE   (SUCCINYL-COA SYNTHETASE, BETA CHAIN) (SCS-BETA) (FRAGMENT).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- CATALYTIC ACTIVITY: SUCCINATE + COA + GTP = SUCCINYL-COA + GDP +
CC       ORTHOPHOSPHATE.
CC   -!- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC       ACID CYCLE.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC       OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
DR   HSC-2DPAGE; P99507; DOG.
DR   INTERPRO; IPR000303; -.
DR   PROSITE; PS01217; SUCCINYL_COA_LIG_3; PARTIAL.
KW   Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2248 MW;  BE8AEFD54DBDAC2E CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 20;
  Best Local Similarity   33.3%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQ 19
           |:: :: ||
Db       5 EYMSMELLQ 13


RESULT  15
DCMS_PSECA
ID   DCMS_PSECA     STANDARD;      PRT;    21 AA.
AC   P19921;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] SMALL CHAIN (EC 1.2.2.4)
DE   (FRAGMENT).
OS   Pseudomonas carboxydovorans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Oligotropha.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=OM5;
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 = CO(2) +
CC       2 H(+) + FERRICYTOCHROME B-561.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0144; PL0144.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2270 MW;  68D43806294O1B9C CRC64;
```

```
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida;
OC   Funariidae; Funariales; Funariaceae; Physcomitrella.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PROTONEMA;
RX   MEDLINE; 97275459.
RA   Kasten B., Buck F., Nuske J., Reski R.;
RT   "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT   plastid enzymes.";
RL   Planta 201:261-272(1997).
CC   -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC       GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC       SUBUNIT.
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   MENDEL; 13023; PHYpa;atpB;1.
DR   INTERPRO; IPR000194; -.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW   ATP synthesis; Chloroplast; Thylakoid membrane; CF(1);
KW   Hydrolase; ATP-binding; Hydrogen ion transport.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2298 MW;  9558E4F5AC89D81A CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFP 10
           :|  ||
Db      16 LDIXFP 21


RESULT  10
DPP4_BOVIN
ID   DPP4_BOVIN     STANDARD;      PRT;    20 AA.
AC   P81425;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   DIPEPTIDYL PEPTIDASE IV (EC 3.4.14.5) (DPP IV) (ADENOSINE DEAMINASE
DE   COMPLEXING PROTEIN) (ADCP-I) (FRAGMENTS).
GN   DPP4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 98293306.
RA   Ben-Shooshan I., Parola A.H.;
RT   "The CP-I subunit of adenosine deaminase complexing protein from calf
RT   kidney is identical to human, mouse, and rat dipeptidyl peptidase
RT   IV.";
RL   Comp. Biochem. Physiol. 119B:289-292(1998).
CC   -!- FUNCTION: REMOVES N-TERMINAL DIPEPTIDES SEQUENTIALLY FROM
CC       POLYPEPTIDES HAVING UNSUBSTITUTED N-TERMINI PROVIDED THAT THE
CC       PENULTIMATE RESIDUE IS PROLINE. BINDS AND REGULATES THE ACTIVITY
CC       OF ADA.
CC   -!- CATALYTIC ACTIVITY: DIPEPTIDYL-POLYPEPTIDE + H(2)O = DIPEPTIDE +
CC       POLYPEPTIDE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS IN
CC       A SOLUBLE FORM (BY SIMILARITY).
CC   -!- PTM: THE SOLUBLE FORM (SDPP) DERIVES FROM THE MEMBRANE FORM (MDPP)
CC       BY PROTEOLYTIC PROCESSING.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE
CC       PROLYL OLIGOPEPTIDASE FAMILY.
DR   INTERPRO; IPR002471; -.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW   Hydrolase; Dipeptidase; Serine protease; Transmembrane; Glycoprotein;
KW   Signal-anchor.
FT   NON_TER      1      1
FT   NON_CONS     10     11
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2356 MW;  28B78E62639F7276 CRC64;


  Query Match             17.6%;  Score 19.5;  DB 1;  Length 20;
  Best Local Similarity   33.3%;  Pred. No. 3.4e+03;
  Matches    5;  Conservative    6;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 LLLQMDFGFPEHLLV 15
           |||::   : |:::|
Db       5 LLLEV---YTENIIV 16


RESULT  11
OXLA_OPHHA
ID   OXLA_OPHHA     STANDARD;      PRT;    19 AA.
AC   P81383;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   L-AMINO ACID OXIDASE (EC 1.4.3.2) (LAO) (LAAO) (FRAGMENT).
OS   Ophiophagus hannah (King cobra) (Naja hannah).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Elapinae; Ophiophagus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 94361525.
RA   Ponnudurai G., Chung M.C.M., Tan N.-H.;
RT   "Purification and properties of the L-amino acid oxidase from Malayan
RT   pit viper (Calloselasma rhodostoma) venom.";
RL   Arch. Biochem. Biophys. 313:373-378(1994).
RN   [2]
RP   SEQUENCE OF 1-15.
RC   TISSUE=VENOM;
RX   MEDLINE; 97449790.
RA   Ahn M.Y., Lee B.M., Kim Y.S.;
RT   "Characterization and cytotoxicity of L-amino acid oxidase from the
RT   venom of king cobra (Ophiophagus hannah).";
RL   Int. J. Biochem. Cell Biol. 29:911-919(1997).
CC   -!- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) = A 2-OXO ACID +
CC       NH(3) + H(2)O(2).
CC   -!- COFACTOR: FAD.
CC   -!- SUBUNIT: HOMODIMER (PROBABLE).
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC       STRONG, TO MOUSE FIG-1.
KW   Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Venom.
FT   CONFLICT     1      1       H -> S (IN REF. 2).
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2298 MW;  DD911A5B414F1427 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+03;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPEH 12
           |:  |  ||:
Db       5 LEESFQEPEY 14


RESULT  12
UP21_UPEIN
```

```
CC       MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC       IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC       REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC       UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR   FLYBASE; FBgn0010375; Mmp-P1.
DR   INTERPRO; IPR001844; -.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW   Chaperone; ATP-binding; Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1243 MW;  78501A366365A6DB CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 11;
  Best Local Similarity   18.2%;  Pred. No. 1.5e+03;
  Matches    2;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPE 11
           ::::  :| |:
Db       1 VIIEQSWGSPK 11


RESULT   6
TEML_RANTE
ID   TEML_RANTE     STANDARD;      PRT;    11 AA.
AC   P56923;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN L.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1194 MW;  1E990549B3372724 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 LLVDFLQSL 21
           || : |:||
Db       1 LLPNLLKSL 9


RESULT   7
DCMM_PSECF
ID   DCMM_PSECF     STANDARD;      PRT;    14 AA.
AC   P19914;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0142; PL0142.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1756 MW;  65583C6D1FB7C25B CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 14;
  Best Local Similarity   15.4%;  Pred. No. 2e+03;
  Matches    2;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPEHL 13
           :: : ::  |:|:
Db       2 MIPRFEYHAPKHV 14


RESULT   8
MDH_SYNY4
ID   MDH_SYNY4      STANDARD;      PRT;    16 AA.
AC   P80460;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Synechocystis sp. (strain PCC 6714).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN   [1]
RP   SEQUENCE.
RA   Naterstad K., Synstad B., Sirevag R.;
RL   Submitted (SEP-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   INTERPRO; IPR001252; -.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1780 MW;  61D1896F14E81984 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LQMDFG 8
           |  |||
Db       9 LTTDFG 14


RESULT   9
ATPB_PHYPA
ID   ATPB_PHYPA     STANDARD;      PRT;    21 AA.
AC   P80658;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (FRAGMENT).
GN   ATPB.
OS   Physcomitrella patens (Moss).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88331073.
RA   Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT   "Isolation and characterization of a laminin-binding protein from rat
RT   and chick muscle.";
RL   J. Cell Biol. 107:687-697(1988).
CC   -!- FUNCTION: CALSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC       CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC       IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH
CC       A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC       TO 50 MOLES OF CALCIUM. ALSO BINDS LAMININ.
CC   -!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE
CC       SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF
CC       FAST SKELETAL MUSCLE CELLS. ASPARTACTIN IS FOUND IN THE BASAL
CC       LAMINA SURROUNDING INDIVIDUAL MUSCLE FIBERS.
CC   -!- TISSUE SPECIFICITY: SKELETAL AND HEART MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
DR   PIR; A31049; A31049.
DR   INTERPRO; IPR001393; -.
DR   PROSITE; PS00864; CALSEQUESTRIN_2; PARTIAL.
DR   PROSITE; PS00863; CALSEQUESTRIN_1; 1.
KW   Muscle; Glycoprotein; Calcium-binding.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2238 MW;  92ADE04FC2A69280 CRC64;


  Query Match             19.8%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 FPEHLLVD 16
           |||:  ||
Db       6 FPEYDGVD 13


RESULT   3
COXN_THUOB
ID   COXN_THUOB     STANDARD;      PRT;    20 AA.
AC   P80980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2303 MW;  0A33BD34006E5AA6 CRC64;


  Query Match             19.8%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 DFGFPEH 12
           | | | |
Db      14 DNGMPVH 20


RESULT   4
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;     9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83160932.
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   INTERPRO; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN       1     >9       PENTAXIN.
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             18.9%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 GFPEHLLV 15
           |||   |:
Db       1 GFPGKSLI 8


RESULT   5
CH60_DROME
ID   CH60_DROME     STANDARD;      PRT;    11 AA.
AC   P35380;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MITOCHONDRIAL MATRIX PROTEIN P1 (60 KDA CHAPERONIN) (HEAT SHOCK
DE   PROTEIN 60) (HSP-60) (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).
GN   MMP-P1 OR HSP60.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VALLECAS; TISSUE=WING IMAGINAL DISK;
RX   MEDLINE; 93272852.
RA   Santaren J.F., van Damme J., Puype M., Vandekerckhove J.,
RA   Garcia-Bellido A.;
RT   "Identification of Drosophila wing imaginal disc proteins by two-
RT   dimensional gel analysis and microsequencing.";
RL   Exp. Cell Res. 206:220-226(1993).
CC   -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:38:16 ; Search time 18.39 Seconds
(without alignments)
38.219 Million cell updates/sec

Title: US-08-934-367-10
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1213

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	21.6	14	1	ANGT_HORSE	P01016 equus cabal
2	22	19.8	20	1	CAQS_RAT	P19633 rattus norv
3	22	19.8	20	1	COXN_THUOB	P80980 thunnus obe
4	21	18.9	9	1	SAMP_MUSCA	P19095 mustelus ca
5	20	18.0	11	1	CH60_DROME	P35380 drosophila
6	20	18.0	11	1	TEML_RANTE	P56923 rana tempor
7	20	18.0	14	1	DCMM_PSECF	P19914 pseudomonas
8	20	18.0	16	1	MDH_SYNY4	P80460 synechocyst
9	20	18.0	21	1	ATPB_PHYPA	P80658 physcomitre
10	19.5	17.6	20	1	DPP4_BOVIN	P81425 bos taurus
11	19	17.1	19	1	OXLA_OPHHA	P81383 ophiophagus
12	19	17.1	19	1	UP21_UPEIN	P82027 uperoleia i
13	19	17.1	19	1	UP25_UPEIN	P82031 uperoleia i
14	19	17.1	20	1	SUCB_CANFA	P99507 canis famil
15	19	17.1	21	1	DCMS_PSECA	P19921 pseudomonas
16	19	17.1	21	1	YD90_HAEIN	P45194 haemophilus
17	18	16.2	7	1	ALL7_CYDPO	P82158 cydia pomon
18	18	16.2	8	1	ALL5_CALVO	P41841 calliphora
19	18	16.2	8	1	ALL5_CYDPO	P82156 cydia pomon
20	18	16.2	12	1	GRAR_RANRU	P40754 rana rugosa
21	18	16.2	13	1	ORCK_ORCLI	P37086 orconectes
22	18	16.2	13	1	PSBP_PINPS	P81668 pinus pinas
23	18	16.2	15	1	LMA2_LOCMI	P38497 locusta mig
24	18	16.2	17	1	FLA2_BARBA	P35634 bartonella
25	18	16.2	18	1	RS4_PSEAE	O52759 pseudomonas
26	18	16.2	19	1	HBB2_UROHA	P18992 uromastyx h
27	18	16.2	19	1	NUO6_SOLTU	P80729 solanum tub
28	18	16.2	20	1	THIO_CANFA	P99505 canis famil
29	18	16.2	20	1	UN05_PINPS	P81674 pinus pinas
30	18	16.2	21	1	CSPS_STRTR	P81622 streptococc
31	18	16.2	21	1	OMP4_PASHA	P80228 pasteurella
32	17	15.3	8	1	CCKN_MACEU	P30369 macropus eu
33	17	15.3	9	1	FAR1_CALVO	P41856 calliphora
34	17	15.3	10	1	CAER_LITXA	P56264 litoria xan
35	17	15.3	12	1	OPS3_DROVI	P17645 drosophila
36	17	15.3	12	1	PSP3_PHYPA	P80662 physcomitre
37	17	15.3	13	1	SODM_CANFA	P54712 canis famil
38	17	15.3	14	1	HCYA_MEGCR	Q10583 megathura c
39	17	15.3	15	1	LPF_ECOLI	P03057 escherichia
40	17	15.3	16	1	LPK1_LOCMI	P20404 locusta mig
41	17	15.3	17	1	GAST_MACMU	P33714 macaca mula
42	17	15.3	19	1	IRBP_CAVPO	P12666 cavia porce
43	17	15.3	19	1	UP24_UPEIN	P82030 uperoleia i
44	17	15.3	20	1	CPA7_PAPSP	P80055 papio sp. (
45	17	15.3	20	1	LPP3_HUMAN	P56643 homo sapien

ALIGNMENTS

RESULT 1
ANGT_HORSE
ID ANGT_HORSE STANDARD; PRT; 14 AA.
AC P01016;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE ANGIOTENSINOGEN (FRAGMENT).
GN AGT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE.
RA Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
RT "The preparation, purification, and amino acid sequence of a
RT polypeptide renin substrate.";
RL J. Exp. Med. 106:439-453(1957).
CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC BALANCE OF BODY FLUIDS.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR PIR; A01250; A01250.
DR INTERPRO; IPR000215; -.
DR PROSITE; PS00284; SERPIN; PARTIAL.
KW Vasoconstrictor; Plasma; Serpin.
FT PEPTIDE 1 10 ANGIOTENSIN I.
FT PEPTIDE 1 8 ANGIOTENSIN II.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1759 MW; 2E9921F8EEEFBDD7 CRC64;

Query Match 21.6%; Score 24; DB 1; Length 14;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 PEHLLV 15
| ||||
Db 7 PFHLLV 12

RESULT 2
CAQS_RAT
ID CAQS_RAT STANDARD; PRT; 20 AA.
AC P19633;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CALSEQUESTRIN, SKELETAL MUSCLE ISOFORM (ASPARTACTIN) (LAMININ-BINDING
DE PROTEIN) (FRAGMENT).
GN CASQ1.
OS Rattus norvegicus (Rat).

```
DR   EMBL; U04282; AAC45541.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1407 MW;  08FAB3930E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


Search completed: December 21, 2000, 08:37:55
Job time: 288 sec
```

```
RA   Menting J.G., Cornish E., Scopes R.K.;
RT   "Purification and partial characterization of NADPH-cytochrome c
RT   reductase from Petunia hybrida flowers.";
RL   Plant Physiol. 106:643-650(1994).
SQ   SEQUENCE   20 AA;  2390 MW;  8BCDA6F8CF7EEDEE CRC64;


  Query Match             22.5%;  Score 25;  DB 10;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 QMDFGFPEHL 13
          :||| : | |
Db      1 RMDFIYEEEL 10


RESULT  12
Q9PS42
ID   Q9PS42      PRELIMINARY;      PRT;    22 AA.
AC   Q9PS42;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PHENOBARBITAL-INDUCED 48 KDA CYTOCHROME P-450 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92406903.
RA   Nakai K., Ward A.M., Gannon M., Rifkind A.B.;
RT   "Beta-naphthoflavone induction of a cytochrome P-450 arachidonic acid
RT   epoxygenase in chick embryo liver distinct from the aryl hydrocarbon
RT   hydroxylase and from phenobarbital-induced arachidonate epoxygenase.";
RL   J. Biol. Chem. 267:19503-19512(1992).
SQ   SEQUENCE   22 AA;  2377 MW;  6BA9C990EE33E9C1 CRC64;


  Query Match             22.1%;  Score 24.5;  DB 13;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 2e+03;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy      5 MDF-GFPEHLLV 15
          ||| | |  ||:
Db      1 MDFLGLPTILLL 12


RESULT  13
O31364
ID   O31364      PRELIMINARY;      PRT;    13 AA.
AC   O31364;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G25;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93700; AAC45534.1; -.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1449 MW;  08FAB9BB0E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LLQMDFG 8
          :| :|||
Db      6 ILVLDFG 12


RESULT  14
O31365
ID   O31365      PRELIMINARY;      PRT;    13 AA.
AC   O31365;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP90;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93701; AAC45536.1; -.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1382 MW;  08EFB4180E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 LLQMDFG 8
          :| :|||
Db      6 ILVLDFG 12


RESULT  15
O34622
ID   O34622      PRELIMINARY;      PRT;    13 AA.
AC   O34622;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT39 (ILL-1), DN127CL9-2, AND 25015;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DN127CL9-2, AND 25015;
RX   MEDLINE; 95154673.
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U93694; AAC45522.1; -.
DR   EMBL; U04280; AAC45543.1; -.
```

```
RESULT   7
Q9ZG55
ID   Q9ZG55      PRELIMINARY;      PRT;    21 AA.
AC   Q9ZG55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ATP-BINDING PROTEIN (FRAGMENT).
GN   RECF.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087306; AAD04082.1; -.
KW   ATP-binding.
FT   NON_TER       1      1
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2336 MW;  0185D9AC428276D9 CRC64;

  Query Match             23.9%;  Score 26.5;  DB 2;  Length 21;
  Best Local Similarity   41.7%;  Pred. No. 9.2e+02;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       2 LLQMDFGFPEHL 13
           :||:   ||:|:
Db       8 ILQL-ISFPKHI 18


RESULT   8
O02830
ID   O02830      PRELIMINARY;      PRT;    22 AA.
AC   O02830;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PRO ALPHA1 TYPE II COLLAGEN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96377339.
RA   Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA   Vuorio E.;
RT   "Evidence for insufficient chondrocytic differentiation during repair
RT   of full-thickness defects of articular cartilage.";
RL   Matrix Biol. 15:39-47(1996).
DR   EMBL; S83370; AAB50773.1; -.
DR   INTERPRO; IPR000885; -.
DR   PFAM; PF01410; COLFI; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   22 AA;  2292 MW;  2E05FE169844236D CRC64;

  Query Match             23.4%;  Score 26;  DB 6;  Length 22;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFPE 11
           || | ||
Db       3 MDIGGPE 9


RESULT   9
O94554
ID   O94554      PRELIMINARY;      PRT;    16 AA.
AC   O94554;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CALMODULIN KINASE 2 (FRAGMENT).
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alemany V., Aligue R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57982; AAD09466.1; -.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1846 MW;  4A673B1F4C3288D9 CRC64;

  Query Match             22.5%;  Score 25;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFP 10
           ||||
Db       5 FGFP 8


RESULT   10
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;

  Query Match             22.5%;  Score 25;  DB 13;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPEHLLVD 16
           ||    | |  | :|
Db       1 LLAQGLGXPTALALD 15


RESULT   11
Q9S878
ID   Q9S878      PRELIMINARY;      PRT;    20 AA.
AC   Q9S878;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4) (FRAGMENT).
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Petunia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95083756.
```

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE; 97312006.
RA   Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT   "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT   sigma 70-like promoter.";
RL   Microbiology 143:1681-1690(1997).
DR   EMBL; U62901; AAB62742.1; -.
FT   NON_TER      21      21
SQ   SEQUENCE   21 AA;  2432 MW;  F33E1EC548BD5B33 CRC64;


  Query Match             27.0%;  Score 30;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           ::| ||| :|
Db       7 IEFDFPEGIL 16


RESULT   3
Q9TWH5
ID   Q9TWH5      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   40 KDA GAP JUNCTION PROTEIN (FRAGMENT).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Heliothis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95347000.
RA   Ryerse J.S.;
RT   "Immunocytochemical, electrophoresis, and immunoblot analysis of
RT   Heliothis virescens gap junctions isolated in the presence and absence
RT   of protease inhibitors.";
RL   Cell Tissue Res. 281:179-186(1995).
SQ   SEQUENCE   20 AA;  2304 MW;  A298D3EB3E89586B CRC64;


  Query Match             26.1%;  Score 29;  DB 5;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 3.5e+02;
  Matches    6;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPEHLLVDF 17
           ::  :| |: | |  ||
Db       3 VIFNIDGGYLEFLTRDF 19


RESULT   4
Q9UGS1
ID   Q9UGS1      PRELIMINARY;      PRT;    12 AA.
AC   Q9UGS1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ796I17.4 (NOVEL PROTEIN SIMILAR TO GS2) (FRAGMENT).
GN   DJ796I17.4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035398; CAB63074.1; -.
FT   NON_TER      1       1
SQ   SEQUENCE   12 AA;  1430 MW;  AF7740ABECB69AA6 CRC64;


  Query Match             24.3%;  Score 27;  DB 4;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGF 9
           ::|:||  |
Db       1 MILEMDMSF 9


RESULT   5
Q9TRQ7
ID   Q9TRQ7      PRELIMINARY;      PRT;    14 AA.
AC   Q9TRQ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN PEPTIDE L-8, CAP-50-ANNEXIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92317074.
RA   Mizutani A., Usuda N., Tokumitsu H., Minami H., Yasui K.,
RA   Kobayashi R., Hidaka H.;
RT   "CAP-50, a newly identified annexin, localizes in nuclei of cultured
RT   fibroblast 3Y1 cells.";
RL   J. Biol. Chem. 267:13498-13504(1992).
SQ   SEQUENCE   14 AA;  1446 MW;  C8322EB96DD9C6C6 CRC64;


  Query Match             24.3%;  Score 27;  DB 6;  Length 14;
  Best Local Similarity   41.7%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFL 18
           ||  |  ::| |
Db       2 FGTDEQAIIDXL 13


RESULT   6
Q9S8E1
ID   Q9S8E1      PRELIMINARY;      PRT;    22 AA.
AC   Q9S8E1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ANNEXIN (FRAGMENT).
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Capsicum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95353219.
RA   Hoshino T., Mizutani A., Chida M., Hidaka H., Mizutani J.;
RT   "Plant annexin form homodimer during Ca(2+)-dependent liposome
RT   aggregation.";
RL   Biochem. Mol. Biol. Int. 35:749-755(1995).
SQ   SEQUENCE   22 AA;  2465 MW;  A36D11A1FE311F0F CRC64;


  Query Match             24.3%;  Score 27;  DB 10;  Length 22;
  Best Local Similarity   41.7%;  Pred. No. 8e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFL 18
           :|  | |::| |
Db       3 WGTDEKLIIDIL 14
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:54 ; Search time 37.32 Seconds
(without alignments)
55.044 Million cell updates/sec

Title: US-08-934-367-10
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4962

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	27.9	19	6	Q9TRR6	Q9trr6 oryctolagus
2	30	27.0	21	2	P70861	P70861 borrelia bu
3	29	26.1	20	5	Q9TWH5	Q9twh5 heliothis v
4	27	24.3	12	4	Q9UGS1	Q9ugs1 homo sapien
5	27	24.3	14	6	Q9TRQ7	Q9trq7 bos taurus
6	27	24.3	22	10	Q9S8E1	Q9s8e1 capsicum an
7	26.5	23.9	21	2	Q9ZG55	Q9zg55 chlamydia t
8	26	23.4	22	6	O02830	O02830 oryctolagus
9	25	22.5	16	3	O94554	O94554 schizosacch
10	25	22.5	19	13	Q9PS70	Q9ps70 gallus gall
11	25	22.5	20	10	Q9S878	Q9s878 petunia hyb
12	24.5	22.1	22	13	Q9PS42	Q9ps42 gallus gall
13	24	21.6	13	2	O31364	O31364 borrelia ga
14	24	21.6	13	2	O31365	O31365 borrelia ga
15	24	21.6	13	2	O34622	O34622 borrelia bu
16	24	21.6	18	1	Q9UYK7	Q9uyk7 pyrococcus
17	24	21.6	19	5	Q9W5Q8	Q9w5q8 drosophila
18	24	21.6	20	6	Q9TTG3	Q9ttg3 ateles belz
19	24	21.6	21	2	Q9X3D0	Q9x3d0 prochloroco
20	23	20.7	8	2	O85406	O85406 coxiella bu
21	23	20.7	9	5	O96417	O96417 drosophila
22	23	20.7	10	2	Q9X534	Q9x534 leclercia a
23	23	20.7	10	2	Q9X533	Q9x533 escherichia
24	23	20.7	13	2	O31296	O31296 borrelia af
25	23	20.7	13	2	O34770	O34770 borrelia af
26	23	20.7	13	11	O35758	O35758 rattus norv
27	23	20.7	17	2	Q9X3I3	Q9x3i3 prochloroco
28	23	20.7	18	13	P82068	P82068 litoria gen
29	23	20.7	19	11	Q9QVE3	Q9qve3 rattus sp.
30	23	20.7	20	13	Q9PRN5	Q9prn5 oncorhynchu
31	23	20.7	21	4	Q16575	Q16575 homo sapien
32	23	20.7	21	8	Q9ZYB7	Q9zyb7 spinaria sp
33	23	20.7	21	10	Q41496	Q41496 solanum tub
34	23	20.7	21	12	Q80817	Q80817 human t-cel
35	23	20.7	22	1	P71537	P71537 methanosarc
36	22.5	20.3	17	6	Q9TR11	Q9tr11 bos taurus
37	22	19.8	13	4	Q9UPE7	Q9upe7 homo sapien
38	22	19.8	14	2	Q56945	Q56945 yersinia ps
39	22	19.8	15	10	Q41526	Q41526 triticum ae
40	22	19.8	17	8	Q9T2S2	Q9t2s2 solanum tub
41	22	19.8	17	11	Q9QUJ4	Q9quj4 mus sp. mep
42	22	19.8	19	9	Q38371	Q38371 bacteriopha
43	22	19.8	20	5	Q9TWI4	Q9twi4 ceratitis c
44	22	19.8	20	6	Q9TRA1	Q9tra1 bos taurus
45	22	19.8	20	13	Q9PRM4	Q9prm4 gallus gall

ALIGNMENTS

RESULT 1
Q9TRR6
ID Q9TRR6 PRELIMINARY; PRT; 19 AA.
AC Q9TRR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CALCYCLIN-ASSOCIATED PROTEIN, CAP50=CA2+/PHOSPHOLIPID-BINDING PROTEIN
DE L-14 FRAGMENT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 92250478.
RA Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT "A calcyclin-associated protein is a newly identified member of the
RT Ca2+/phospholipid-binding proteins, annexin family.";
RL J. Biol. Chem. 267:8919-8924(1992).
SQ SEQUENCE 19 AA; 2018 MW; 9A54062504B8322E CRC64;

Query Match 27.9%; Score 31; DB 6; Length 19;
Best Local Similarity 43.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```
Qy       7 FGFPEHLLVDFLQSLS 22
           ||  |  ::| | | |
Db       2 FGTDEQAIIDXLGSRS 17
```

RESULT 2
P70861
ID P70861 PRELIMINARY; PRT; 21 AA.
AC P70861;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE THDF (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]

```
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45043; AAB04194.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  14
Q73479
ID   Q73479      PRELIMINARY;      PRT;    18 AA.
AC   Q73479;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45045; AAB04196.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  15
Q73481
ID   Q73481      PRELIMINARY;      PRT;    18 AA.
AC   Q73481;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45047; AAB04198.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


Search completed: December 21, 2000, 08:35:39
Job time: 602 sec
```

```
  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  10
Q73471
ID   Q73471      PRELIMINARY;      PRT;    18 AA.
AC   Q73471;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45037; AAB04188.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  11
Q73473
ID   Q73473      PRELIMINARY;      PRT;    18 AA.
AC   Q73473;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45039; AAB04190.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  12
Q73475
ID   Q73475      PRELIMINARY;      PRT;    18 AA.
AC   Q73475;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45041; AAB04192.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  13
Q73477
ID   Q73477      PRELIMINARY;      PRT;    18 AA.
AC   Q73477;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45029; AAB04180.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT   7
Q73465
ID   Q73465      PRELIMINARY;      PRT;    18 AA.
AC   Q73465;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45031; AAB04182.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT   8
Q73467
ID   Q73467      PRELIMINARY;      PRT;    18 AA.
AC   Q73467;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45033; AAB04184.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT   9
Q73469
ID   Q73469      PRELIMINARY;      PRT;    18 AA.
AC   Q73469;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45035; AAB04186.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;
```

```
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-13(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45053; AAB04204.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2191 MW;  B36910F4583BA0C3 CRC64;


  Query Match             29.9%;  Score 32;  DB 12;  Length 18;
  Best Local Similarity   38.5%;  Pred. No. 63;
  Matches    5;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           ::|: |  ||::|
Db       2 LLKWGFTTPDKKH 14


RESULT   3
Q9W9U1
ID   Q9W9U1      PRELIMINARY;      PRT;    19 AA.
AC   Q9W9U1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   GLYCOPROTEIN G (FRAGMENT).
GN   US4.
OS   Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=20620, 9400371, 7709642, 7709609;
RA   Eberle R., Smith A., Black D.;
RT   "Molecular evidence for distinct genotypes of monkey B virus
RT   (Herpesvirus simiae) which are related to the Macaque host species.";
RL   J. Virol. 0:0-0(1998).
DR   EMBL; AF082804; AAC34447.1; -.
DR   EMBL; AF082812; AAC34111.1; -.
DR   EMBL; AF082813; AAC34114.1; -.
DR   EMBL; AF082814; AAC34117.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2220 MW;  E585178F064B278B CRC64;


  Query Match             28.0%;  Score 30;  DB 12;  Length 19;
  Best Local Similarity   54.5%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
           |   || | ||
Db       2 RSPSQHRVRYT 12


RESULT   4
Q9WLD2
ID   Q9WLD2      PRELIMINARY;      PRT;    20 AA.
AC   Q9WLD2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   GLYCOPROTEIN G (FRAGMENT).
GN   US4.
OS   Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=12930;
RA   Eberle R., Smith A., Black D.;
RT   "Molecular evidence for distinct genotypes of monkey B virus
RT   (Herpesvirus simiae) which are related to the Macaque host species.";
RL   J. Virol. 0:0-0(1998).
DR   EMBL; AF082808; AAC34459.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2377 MW;  E58517E7DBAB278B CRC64;


  Query Match             28.0%;  Score 30;  DB 12;  Length 20;
  Best Local Similarity   54.5%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
           |   || | ||
Db       3 RSPSQHRVRYT 13


RESULT   5
Q73461
ID   Q73461      PRELIMINARY;      PRT;    18 AA.
AC   Q73461;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45027; AAB04178.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36918D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT   6
Q73463
ID   Q73463      PRELIMINARY;      PRT;    18 AA.
AC   Q73463;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:38 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-9
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	31.8	17	3	Q06800	Q06800 saccharomyc
2	32	29.9	18	12	Q73487	Q73487 human immun
3	30	28.0	19	12	Q9W9U1	Q9w9u1 simian herp
4	30	28.0	20	12	Q9WLD2	Q9wld2 simian herp
5	29	27.1	18	12	Q73461	Q73461 human immun
6	29	27.1	18	12	Q73463	Q73463 human immun
7	29	27.1	18	12	Q73465	Q73465 human immun
8	29	27.1	18	12	Q73467	Q73467 human immun
9	29	27.1	18	12	Q73469	Q73469 human immun
10	29	27.1	18	12	Q73471	Q73471 human immun
11	29	27.1	18	12	Q73473	Q73473 human immun
12	29	27.1	18	12	Q73475	Q73475 human immun
13	29	27.1	18	12	Q73477	Q73477 human immun
14	29	27.1	18	12	Q73479	Q73479 human immun
15	29	27.1	18	12	Q73481	Q73481 human immun
16	29	27.1	18	12	Q73483	Q73483 human immun
17	29	27.1	18	12	Q73485	Q73485 human immun
18	29	27.1	18	12	Q73489	Q73489 human immun
19	29	27.1	18	12	Q73491	Q73491 human immun
20	29	27.1	18	12	Q73493	Q73493 human immun
21	29	27.1	18	12	Q73495	Q73495 human immun
22	29	27.1	18	12	Q73497	Q73497 human immun
23	29	27.1	18	12	Q73499	Q73499 human immun
24	29	27.1	18	12	Q73501	Q73501 human immun
25	29	27.1	18	12	Q73503	Q73503 human immun
26	29	27.1	18	12	Q73505	Q73505 human immun
27	29	27.1	18	12	Q73507	Q73507 human immun
28	29	27.1	18	12	Q73509	Q73509 human immun
29	29	27.1	18	12	Q73511	Q73511 human immun
30	29	27.1	18	12	Q73513	Q73513 human immun
31	29	27.1	18	12	Q73515	Q73515 human immun
32	29	27.1	18	12	Q73519	Q73519 human immun
33	29	27.1	18	12	Q73521	Q73521 human immun
34	29	27.1	18	12	Q73525	Q73525 human immun
35	29	27.1	18	12	Q73527	Q73527 human immun
36	29	27.1	18	12	Q73529	Q73529 human immun
37	28	26.2	14	12	Q89818	Q89818 murine minu
38	28	26.2	16	12	Q79458	Q79458 human immun
39	28	26.2	19	12	O90625	O90625 simian herp
40	27	25.2	17	12	Q73531	Q73531 human immun
41	27	25.2	17	12	Q73533	Q73533 human immun
42	27	25.2	17	12	Q73535	Q73535 human immun
43	27	25.2	17	12	Q73537	Q73537 human immun
44	27	25.2	17	12	Q73539	Q73539 human immun
45	27	25.2	17	12	Q73541	Q73541 human immun

ALIGNMENTS

RESULT 1
Q06800
ID Q06800 PRELIMINARY; PRT; 17 AA.
AC Q06800;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE N1780.
GN N1780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96287653.
RA Nasr F., Becam A.M., Herbert C.J.;
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
RT 24 complete open reading frames: 18 correspond to new genes, one of
RT which encodes a protein similar to the human myotonic dystrophy
RT kinase.";
RL Yeast 12:169-175(1996).
DR EMBL; X92517; CAA63292.1; -.
SQ SEQUENCE 17 AA; 2139 MW; BD7E9AAFADF754AF CRC64;

Query Match 31.8%; Score 34; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      12 QQHSVAYT 19
           ||||| :|
Db       3 QQHSVCFT 10
```

RESULT 2
Q73487
ID Q73487 PRELIMINARY; PRT; 18 AA.
AC Q73487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE REVERSE TRANSCRIPTASE (FRAGMENT).

```
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A35704; A35704.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum; Olfaction.
FT   NON_TER       1      1
FT   VARIANT       6      6       G -> D.
FT   VARIANT      11     11       A -> E.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2058 MW;  F80746F76CCD77FF CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      8 PRPDQQ 13
          | | ||
Db      5 PGPQQQ 10


RESULT  13
UPA3_HUMAN
ID   UPA3_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30089;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 11) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PLASMA;
RX   MEDLINE; 93092937.
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 4.6, ITS MW IS: 46 KDA.
DR   SWISS-2DPAGE; P30089; HUMAN.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1056 MW;  26F2B1BAF769C737 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 LFPRPD 11
          |||  |
Db      3 LFPXTD 8


RESULT  14
TKNA_HORSE
ID   TKNA_HORSE     STANDARD;      PRT;    11 AA.
AC   P01290;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
GN   TAC1 OR NKNA OR TAC2 OR NKA.
OS   Equus caballus (Horse), and Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=HORSE;
RA   Studer R.O., Trzeciak A., Lergier W.;
RT   "Isolation and amino-acid sequence of substance P from horse
RT   intestine.";
RL   Helv. Chim. Acta 56:860-866(1973).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=C.PORCELLUS;
RX   MEDLINE; 90044685.
RA   Murphy R.;
RT   "Primary amino acid sequence of guinea-pig substance P.";
RL   Neuropeptides 14:105-110(1989).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; A01558; SPHO.
DR   PIR; A60654; A60654.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1349 MW;  3E757FE3C9D6C6C7 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 6.8e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          |:| |
Db      2 PKPQQ 6


RESULT  15
ECDE_LYMDI
ID   ECDE_LYMDI     STANDARD;      PRT;    13 AA.
AC   P80941;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   TESTIS ECDYSIOTROPIN PEPTIDE E (TE).
OS   Lymantria dispar (Gypsy moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Lymantriidae; Lymantria.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 97387807.
RA   Loeb M.J., Wagner R.M., Woods C.W., Gelman D.G., Harrison D.,
RA   Bell R.A.;
RT   "Naturally occurring analogs of Lymantria testis ecdysiotropin, a
RT   gonadotropin isolated from brains of Lymantria dispar pupae.";
RL   Arch. Insect Biochem. Physiol. 36:37-50(1997).
CC   -!- FUNCTION: STIMULATES SYNTHESIS OF ECDYSTEROID IN THE TESTES
CC       OF LARVAE AND PUPAE.
SQ   SEQUENCE   13 AA;  1357 MW;  1841B4CA3275B764 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      8 PRPDQQHS 15
          | ||   |
Db      5 PNPDTPDS 12



Search completed: December 21, 2000, 08:32:59
Job time: 442 sec
```

```
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; B60409; B60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1269 MW;  3DBA7C37C9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PHPDE 6


RESULT  10
TKN2_PSEGU
ID   TKN2_PSEGU     STANDARD;      PRT;    11 AA.
AC   P42987;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   KASSININ-LIKE PEPTIDE K-II (PG-KII).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; C60409; C60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1246 MW;  3A247C37C9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PNPDE 6


RESULT  11
TKN4_PSEGU
ID   TKN4_PSEGU     STANDARD;      PRT;    11 AA.
AC   P42989;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P-LIKE PEPTIDE I (PG-SPI).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; E60409; E60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1294 MW;  3A247C2CC9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PNPDE 6


RESULT  12
CPAX_BOVIN
ID   CPAX_BOVIN     STANDARD;      PRT;    18 AA.
AC   P22779;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CYTOCHROME P450 2A (OLF2) (OLFACTIVE) (P52) (EC 1.14.14.1) (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 91027757.
RA   Lazard D., Tal N., Rubinstein M., Khen M., Lancet D., Zupko K.;
RT   "Identification and biochemical analysis of novel olfactory-specific
RT   cytochrome P-450IIA and UDP-glucuronosyl transferase.";
RL   Biochemistry 29:7433-7440(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
```

```
  Best Local Similarity   57.1%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 QQHSVAY 18
           ::|||||:
Db       3 RRHSVAW 9


RESULT   6
HELT_HELHO
ID   HELT_HELHO     STANDARD;      PRT;     20 AA.
AC   P46693;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   HELOTHERMINE (FRAGMENT).
OS   Heloderma horridum horridum (Mexican beaded lizard).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 90260878.
RA   Mochca-Morales J., Martin B.M., Possani L.D.;
RT   "Isolation and characterization of helothermine, a novel toxin from
RT   Heloderma horridum horridum (Mexican beaded lizard) venom.";
RL   Toxicon 28:299-309(1990).
CC   -!- FUNCTION: TOXIC TO MICE; INDUCES LETHARGY, PARTIAL PARALYSIS OF
CC       REAR LIMBS AND LOWERING OF BODY TEMPERATURE, SUGGESTING THAT IT
CC       MIGHT BE A HYPOTHERMIC TOXIN.
CC   -!- MISCELLANEOUS: THE COMPLETE PROTEIN HAS AN APPARENT MW OF 25 KDA
CC       AND A PI OF 6.8.
DR   PIR; A34859; A34859.
KW   Toxin.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2156 MW;  91D62B36F7B4F940 CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 4e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 PDQQHSV 16
           ||||  :
Db      14 PDQQTEI 20


RESULT   7
TKNA_ONCMY
ID   TKNA_ONCMY     STANDARD;      PRT;     11 AA.
AC   P28499;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 92298992.
RA   Jensen J., Conlon J.M.;
RT   "Substance-P-related and neurokinin-A-related peptides from the brain
RT   of the cod and trout.";
RL   Eur. J. Biochem. 206:659-664(1992).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S23307; S23307.
DR   PIR; S23308; S23308.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION (BY SIMILARITY).
SQ   SEQUENCE   11 AA;  1358 MW;  214860DEC9D6D1F7 CRC64;


  Query Match             21.5%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
            ||| |
Db       2 PRPHQ 6


RESULT   8
TKNA_SCYCA
ID   TKNA_SCYCA     STANDARD;      PRT;     11 AA.
AC   P41333;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93292508.
RA   Waugh D., Wang Y., Hazon N., Balment R.J., Conlon J.M.;
RT   "Primary structures and biological activities of substance-P-related
RT   peptides from the brain of the dogfish, Scyliorhinus canicula.";
RL   Eur. J. Biochem. 214:469-474(1993).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S33300; S33300.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1278 MW;  214860DEC9D6D867 CRC64;


  Query Match             21.5%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
            ||| |
Db       2 PRPGQ 6


RESULT   9
TKN1_PSEGU
ID   TKN1_PSEGU     STANDARD;      PRT;     11 AA.
AC   P42986;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   KASSININ-LIKE PEPTIDE K-I (PG-KI).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
```

```
RT   "Purification and partial amino acid sequences of two distinct
RT   albumins from turtle plasma.";
RL   Comp. Biochem. Physiol. 118B:367-374(1997).
CC   -!- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC       BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC       BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC       COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- MISCELLANEOUS: IN THE RED-EARED SLIDER TURTLE, THERE ARE TWO FORMS
CC       OF ALBUMIN, ALB-1 AND ALB-2.
CC   -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
DR   INTERPRO; IPR000264; -.
DR   PROSITE; PS00212; ALBUMIN; PARTIAL.
KW   Plasma; Metal-binding; Lipid-binding; Albumin.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1733 MW;  4B7422B89FF73223 CRC64;


  Query Match             23.4%;  Score 25;  DB 1;  Length 15;
  Best Local Similarity   30.0%;  Pred. No. 1.9e+02;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     11 DQQHSVAYTF 20
          |: |:: : |
Db      3 DETHTLGHXF 12


RESULT   3
TKNA_CHICK
ID   TKNA_CHICK     STANDARD;      PRT;    11 AA.
AC   P19850;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=INTESTINE;
RX   MEDLINE; 88204263.
RA   Conlon J.M., Katsoulis S., Schmidt W.E., Thim L.;
RT   "[Arg3]substance P and neurokinin A from chicken small intestine.";
RL   Regul. Pept. 20:171-180(1988).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; JN0023; JN0023.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1377 MW;  21487FE3C9D6C6C7 CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
            ||| |
Db      2 PRPQQ 6


RESULT   4
TKNA_GADMO
ID   TKNA_GADMO     STANDARD;      PRT;    11 AA.
AC   P28498;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Gadus morhua (Atlantic cod).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidei; Gadidae;
OC   Gadus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 92298992.
RA   Jensen J., Conlon J.M.;
RT   "Substance-P-related and neurokinin-A-related peptides from the brain
RT   of the cod and trout.";
RL   Eur. J. Biochem. 206:659-664(1992).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION (BY SIMILARITY).
SQ   SEQUENCE   11 AA;  1315 MW;  214860D759D6C6C7 CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
            ||| |
Db      2 PRPQQ 6


RESULT   5
GALS_SALTY
ID   GALS_SALTY     STANDARD;      PRT;    18 AA.
AC   P41030;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   MGL REPRESSOR AND GALACTOSE ULTRAINDUCTION FACTOR (FRAGMENT).
GN   GALS.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE; 89112167.
RA   Benner-Luger D., Boos W.;
RT   "The mglB sequence of Salmonella typhimurium LT2; promoter analysis
RT   by gene fusions and evidence for a divergently oriented gene coding
RT   for the mgl repressor.";
RL   Mol. Gen. Genet. 214:579-587(1988).
CC   -!- FUNCTION: REPRESSOR OF THE MGL OPERON. BINDS GALACTOSE AND
CC       D-FUCOSE AS INDUCERS. GALS BINDS TO AN OPERATOR DNA SEQUENCE
CC       WITHIN ITS OWN CODING SEQUENCE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LACI FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   STYGENE; SG10521; GALS.
DR   INTERPRO; IPR000843; -.
DR   PROSITE; PS00356; HTH_LACI_FAMILY; PARTIAL.
KW   Transcription regulation; DNA-binding; Repressor.
FT   NON_TER      1      1
SQ   SEQUENCE   18 AA;  2133 MW;  74A77150FA49FB8B CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 18;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:58 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-9
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	25.2	12	1	TKN1_KASMA	P08613 kassina mac
2	25	23.4	15	1	ALB2_TRASC	P81189 trachemys s
3	24	22.4	11	1	TKNA_CHICK	P19850 gallus gall
4	24	22.4	11	1	TKNA_GADMO	P28498 gadus morhu
5	24	22.4	18	1	GALS_SALTY	P41030 salmonella
6	24	22.4	20	1	HELT_HELHO	P46693 heloderma h
7	23	21.5	11	1	TKNA_ONCMY	P28499 oncorhynchu
8	23	21.5	11	1	TKNA_SCYCA	P41333 scyliorhinu
9	22	20.6	11	1	TKN1_PSEGU	P42986 pseudophryn
10	22	20.6	11	1	TKN2_PSEGU	P42987 pseudophryn
11	22	20.6	11	1	TKN4_PSEGU	P42989 pseudophryn
12	22	20.6	18	1	CPAX_BOVIN	P22779 bos taurus
13	21	19.6	9	1	UPA3_HUMAN	P30089 homo sapien
14	21	19.6	11	1	TKNA_HORSE	P01290 equus cabal
15	21	19.6	13	1	ECDE_LYMDI	P80941 lymantria d
16	21	19.6	15	1	MK2A_PALPR	P80409 palomena pr
17	21	19.6	16	1	MK3_PALPR	P80411 palomena pr
18	21	19.6	17	1	PSBL_SYNVU	P12241 synechococc
19	20	18.7	9	1	FAR9_ASCSU	P43172 ascaris suu
20	20	18.7	10	1	BPP2_BOTJA	P01022 bothrops ja
21	20	18.7	11	1	TKNA_RANCA	P22688 rana catesb
22	20	18.7	13	1	BPP1_BOTJA	P01020 bothrops ja
23	20	18.7	13	1	GER1_HORVU	P28525 hordeum vul
24	20	18.7	17	1	SRY_URSAR	P36396 ursus arcto
25	20	18.7	19	1	OXLA_OPHHA	P81383 ophiophagus
26	20	18.7	20	1	ACPH_BOVIN	P80227 bos taurus
27	20	18.7	20	1	LEC3_ARTIN	P18673 artocarpus
28	19	17.8	5	1	PAP2_PARMA	P81864 pardachirus
29	19	17.8	10	1	FARP_LOCMI	P38553 locusta mig
30	19	17.8	11	1	BPP3_BOTIN	P30423 bothrops in
31	19	17.8	11	1	BPP4_BOTIN	P30424 bothrops in
32	19	17.8	11	1	BPPB_AGKHA	P01021 agkistrodon
33	19	17.8	12	1	TKN2_KASMA	P08614 kassina mac
34	19	17.8	12	1	TKN_KASSE	P08611 kassina sen
35	19	17.8	13	1	GER2_HORVU	P28526 hordeum vul
36	19	17.8	13	1	YCIA_SALTY	P25944 salmonella
37	19	17.8	15	1	MK1_PALPR	P80408 palomena pr
38	19	17.8	16	1	MK2B_PALPR	P80410 palomena pr
39	19	17.8	17	1	APID_BOMPA	P81464 bombus pasc
40	19	17.8	17	1	YALA_TRYBB	P17961 trypanosoma
41	19	17.8	20	1	ALAT_PIG	P13191 sus scrofa
42	19	17.8	20	1	PYRR_PYRAP	P37362 pyrrhocoris
43	18.5	17.3	18	1	HEMH_THETS	P80155 theromyzon
44	18	16.8	10	1	UPA2_HUMAN	P30088 homo sapien
45	18	16.8	13	1	CRBL_ICASP	P17237 icaria sp.

ALIGNMENTS

RESULT 1
TKN1_KASMA
ID TKN1_KASMA STANDARD; PRT; 12 AA.
AC P08613;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYLAMBATES KASSININ (GLU(2)-PRO(5) KASSININ).
OS Kassina maculata (African rhacophorid frog) (Hylambates maculatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;
OC Kassina.
RN [1]
RP SEQUENCE.
RA Yasuhara T., Nakajima T., Erspamer G.F., Erspamer V.;
RT "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and
RT hylambatin, in the skin of the African rhacophorid frog Hylambates
RT maculatus.";
RL Biomed. Res. 2:613-617(1981).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR PIR; S10059; S10059.
DR INTERPRO; IPR002040; -.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1376 MW; 3E756D279DD6DAB7 CRC64;

Query Match 25.2%; Score 27; DB 1; Length 12;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PRPDQ 12
|:|||
Db 3 PKPDQ 7

RESULT 2
ALB2_TRASC
ID ALB2_TRASC STANDARD; PRT; 15 AA.
AC P81189;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 68 KDA SERUM ALBUMIN (ALB-2) (FRAGMENT).
OS Trachemys scripta (Red-eared slider turtle) (Pseudemys scripta).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
RN [1]
RP SEQUENCE.
RX MEDLINE; 98103404.
RA Brown M.A., Chambers G.K., Licht P.;

```
  Query Match             22.4%;  Score 24;  DB 2;  Length 19;
  Best Local Similarity   71.4%;  Pred. No. 7.6e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 QQHSVAY 18
           || |:||
Db      12 QQWSMAY 18


RESULT  14
S69166
ferredoxin b - Japanese radish (fragments)
C;Species: Kaiware daikon (Japanese radish)
C;Date: 10-Mar-1998 #sequence_revision 10-Mar-1998 #text_change 17-Apr-1998
C;Accession: S69166
R;Obata, S.; Nishimura, M.; Nagai, K.; Sakihama, N.; Shin, M.
Arch. Biochem. Biophys. 316, 797-802, 1995
A;Title: Four ferredoxins from Japanese radish leaves.
A;Reference number: S69164; MUID:95168867
A;Accession: S69166
A;Molecule type: protein
A;Residues: 1-19 <OBA>
C;Keywords: 2Fe-2S; electron transfer; iron-sulfur protein


  Query Match             22.4%;  Score 24;  DB 2;  Length 19;
  Best Local Similarity   28.6%;  Pred. No. 7.6e+02;
  Matches    4;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 VKFLFPRPDQQHSV 16
           |||: |   ::  :
Db       5 VKFITPETHKEEDM 18


RESULT  15
JP0055
ribosomal protein L30 - Bacillus polymyxa (fragment)
C;Species: Bacillus polymyxa
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994
C;Accession: JP0055
R;Ochi, K.
submitted to JIPID, February 1994
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal pr
A;Reference number: JP0042
A;Accession: JP0055
A;Molecule type: protein
A;Residues: 1-20 <OCH>
C;Keywords: protein biosynthesis; ribosome


  Query Match             22.4%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   35.7%;  Pred. No. 8.1e+02;
  Matches    5;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQHS 15
           :|: |  ||  | :
Db       7 LVRSLIGRPGTQRT 20


Search completed: December 21, 2000, 08:30:06
Job time: 270 sec
```

A;Residues: 1-11 <JEN>
A;Experimental source: brain
C;Function:
A;Description: may play a physiological role in the regulation of cardiovascular and gas
A;Note: substance P is derived by post-translational processing of preprotachykinin A
C;Keywords: neuropeptide; amidated carboxyl end; tachykinin
F;11/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 22.4%; Score 24; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 8 PRPDQ 12
||| |
Db 2 PRPQQ 6

RESULT 9
S47395
T-cell antigen receptor VJ junction beta chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47395
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c
A;Reference number: S47355
A;Accession: S47395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12 <LEH>
A;Cross-references: EMBL:Z35715; NID:g527525; PIDN:CAA84784.1; PID:g527526
C;Keywords: T-cell receptor

Query Match 22.4%; Score 24; DB 2; Length 12;
Best Local Similarity 62.5%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 13 QHSVAYTF 20
| | |||
Db 5 QGSYGYTF 12

RESULT 10
S29207
avenin gamma-4 - oat (fragment)
N;Alternate names: CIP-1; coeliac immunoreactive protein 1
C;Species: Avena sativa (oat)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C;Accession: S29207
R;Rocher, A.; Colilla, F.; Ortiz, M.L.; Mendez, E.
FEBS Lett. 310, 37-40, 1992
A;Title: Identification of the three major coeliac immunoreactive proteins and one alpha
A;Reference number: S29207; MUID:92405739
A;Accession: S29207
A;Molecule type: protein
A;Residues: 1-15 <ROC>
A;Experimental source: endosperm
C;Superfamily: gliadin
C;Keywords: prolamin; seed

Query Match 22.4%; Score 24; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 PRPDQQ 13
| |:||
Db 7 PYPEQQ 12

RESULT 11
A28965
ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain - spinach (fragments)
C;Species: Spinacia oleracea (spinach)
C;Date: 22-Dec-1988 #sequence_revision 22-Dec-1988 #text_change 23-Feb-1997
C;Accession: A28965
R;Mulligan, R.M.; Houtz, R.L.; Tolbert, N.E.
Proc. Natl. Acad. Sci. U.S.A. 85, 1513-1517, 1988
A;Title: Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/
cetylated proline.
A;Reference number: A28965; MUID:88144466
A;Accession: A28965
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <MUL>
C;Keywords: carbon-carbon lyase; carboxy-lyase; chloroplast

Query Match 22.4%; Score 24; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 KFLFPRPD 11
|| || |
Db 6 KFEFPAMD 13

RESULT 12
B25348
UDPglucose--glycogen glucosyltransferase (EC 2.4.1.11) P-2 peptide - rabbit (fragment
N;Alternate names: UDPglucose-glycogen glucosyltransferase
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 06-Dec-1996
C;Accession: B25348
R;Wang, Y.; Bell, A.W.; Hermodson, M.A.; Roach, P.J.
J. Biol. Chem. 261, 16909-16915, 1986
A;Title: Liver isozyme of rabbit glycogen synthase. Amino acid sequences surrounding
A;Reference number: A92570; MUID:87057401
A;Accession: B25348
A;Molecule type: protein
A;Residues: 1-17 <WAN>
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of glycogen by UDPglucose produc
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 22.4%; Score 24; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LFPRP 10
::|||
Db 1 MYPRP 5

RESULT 13
S19613
globin - polychaete (Eudistylia vancouveri) (fragment)
C;Species: Eudistylia vancouveri
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C;Accession: S19613
R;Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp,
J. Mol. Biol. 222, 1109-1129, 1991
A;Title: Hierarchy of globin complexes. The quaternary structure of the extracellular
A;Reference number: S19532; MUID:92106333
A;Accession: S19613
A;Molecule type: protein
A;Residues: 1-19 <QAB>
A;Experimental source: plume
C;Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dod
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

```
RESULT   3
A54083
p190/210, fatty acid synthase, p140exo2 strand exchange protein activator - fission yeas
C;Species: Schizosaccharomyces pombe
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C;Accession: A54083
R;Kaslan, E.; Heyer, W.D.
J. Biol. Chem. 269, 14103-14110, 1994
A;Title: Schizosaccharomyces pombe fatty acid synthase mediates DNA strand exchange in v
A;Reference number: A54083; MUID:94245730
A;Accession: A54083
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <KAS>
A;Note: sequence extracted from NCBI backbone (NCBIP:148744)
C;Superfamily: yeast fatty-acid synthase


  Query Match             24.3%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   36.4%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
           ||: :  :|:|
Db       2 RPEVEQELAHT 12


RESULT   4
A60770
cell surface alloantigen gp60 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-Feb-1997
C;Accession: A60770
R;Schadendorf, D.; Yamaguchi, H.; Old, L.J.; Srivastava, P.K.
J. Immunol. 142, 1621-1625, 1989
A;Title: A novel heteromorphic human cell surface alloantigen, gp60, defined by a human
A;Reference number: A60770; MUID:89140352
A;Accession: A60770
A;Molecule type: protein
A;Residues: 1-14 <SCH>
C;Comment: This protein is an alloantigen in human populations but is not found on cells
C;Comment: This protein exists in both membrane bound and cytosolic forms.
C;Keywords: glycoprotein; polymorphism; surface antigen


  Query Match             23.4%;  Score 25;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 HSVAYT 19
           |||::|
Db       3 HSVSFT 8


RESULT   5
PS0221
gastrin-releasing peptide - laughing frog (fragment)
C;Species: Rana ridibunda (laughing frog)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 03-Mar-1995
C;Accession: PS0221
R;Conlon, J.M.; O'Harte, F.; Vaudry, H.
Biochem. Biophys. Res. Commun. 178, 526-530, 1991
A;Title: Primary structures of the bombesin-like neuropeptides in frog brain show that b
A;Reference number: PQ0177; MUID:91315477
A;Accession: PS0221
A;Molecule type: protein
A;Residues: 1-15 <CON>
A;Experimental source: brain
C;Keywords: neuropeptide


  Query Match             23.4%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PDQQH 14
           | |||
Db       2 PSQQH 6


RESULT   6
A41436
alpha-macroglobulin - green seaturtle (fragment)
C;Species: Chelonia mydas (green seaturtle)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 18-Jun-1993
C;Accession: A41436
R;Osada, T.; Sasaki, T.; Ikai, A.
J. Biochem. 103, 212-217, 1988
A;Title: Purification and characterization of alpha-macroglobulin and ovomacroglobuli
A;Reference number: A41436; MUID:88227890
A;Accession: A41436
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <OSA>


  Query Match             23.4%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 4e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VMVKFLF 7
           |:| |||
Db       8 VLVPFLF 14


RESULT   7
JN0023
substance P - chicken
C;Species: Gallus gallus (chicken)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jul-1997
C;Accession: JN0023
R;Conlon, J.M.; Katsoulis, S.; Schmidt, W.E.; Thim, L.
Regul. Pept. 20, 171-180, 1988
A;Title: [Arg3]substance P and neurokinin A from chicken small intestine.
A;Reference number: JN0023; MUID:88204263
A;Accession: JN0023
A;Molecule type: protein
A;Residues: 1-11 <CON>
C;Superfamily: substance P precursor
C;Keywords: amidated carboxyl end; tachykinin
F;11/Modified site: amidated carboxyl end (Met) #status predicted


  Query Match             22.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
           ||| |
Db       2 PRPQQ 6


RESULT   8
S23306
substance P - Atlantic cod
C;Species: Gadus morhua (Atlantic cod)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Feb-1999
C;Accession: S23306
R;Jensen, J.; Conlon, J.M.
Eur. J. Biochem. 206, 659-664, 1992
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the
A;Reference number: S23186; MUID:92298992
A;Accession: S23306
A;Molecule type: protein
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:06 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-9
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	25.2	12	2	S10059	tachykinin - Afric
2	27	25.2	20	2	JP0053	ribosomal protein
3	26	24.3	20	2	A54083	p190/210, fatty ac
4	25	23.4	14	2	A60770	cell surface alloa
5	25	23.4	15	2	PS0221	gastrin-releasing
6	25	23.4	15	2	A41436	alpha-macroglobuli
7	24	22.4	11	2	JN0023	substance P - chic
8	24	22.4	11	2	S23306	substance P - Atla
9	24	22.4	12	2	S47395	T-cell antigen rec
10	24	22.4	15	2	S29207	avenin gamma-4 - o
11	24	22.4	15	2	A28965	ribulose-bisphosph
12	24	22.4	17	2	B25348	UDPglucose--glycog
13	24	22.4	19	2	S19613	globin - polychaet
14	24	22.4	19	2	S69166	ferredoxin b - Jap
15	24	22.4	20	2	JP0055	ribosomal protein
16	24	22.4	20	2	A34859	helothermine - Mex
17	23	21.5	9	2	A60108	exotoxin A - Strep
18	23	21.5	10	2	S19296	16K protein - poul
19	23	21.5	11	2	S78026	ribosomal protein
20	23	21.5	11	2	S33300	probable substance
21	23	21.5	11	2	S23308	substance P - rain
22	23	21.5	16	2	A60551	leukocyte elastase
23	23	21.5	20	2	A47105	dystroglycan - chi
24	22	20.6	7	4	I56695	hypothetical L2 pr
25	22	20.6	11	2	B60409	kassinin-like pept
26	22	20.6	11	2	C60409	kassinin-like pept
27	22	20.6	11	2	E60409	substance P-like p
28	22	20.6	14	2	S60353	amylopullulanase -
29	22	20.6	18	2	A35704	cytochrome P450 ol
30	22	20.6	18	2	A42016	mammary-derived gr
31	22	20.6	19	2	JP0054	ribosomal protein
32	21	19.6	9	4	S15594	orf 1 rara 5'-regi
33	21	19.6	11	1	SPHO	substance P - hors
34	21	19.6	11	1	A60654	substance P - guin
35	21	19.6	12	2	S47391	T-cell antigen rec
36	21	19.6	15	2	S26515	T-cell receptor al
37	21	19.6	16	2	S57517	T cell receptor be
38	21	19.6	17	2	JH0785	DNA-directed RNA p
39	21	19.6	17	2	S05033	photosystem II pro
40	21	19.6	18	2	S55002	protein 1 - Legion
41	21	19.6	20	2	S18582	hypothetical prote
42	21	19.6	20	2	A56046	urinary tract ston
43	20	18.7	8	2	S66646	cardioacceleratory
44	20	18.7	10	1	XAVI6B	angiotensin-conver
45	20	18.7	10	2	A61007	hementin (EC 3.4.-

ALIGNMENTS

RESULT 1
S10059
tachykinin - African tree frog (Kassina maculata)
N;Alternate names: hylambates-kassinin
C;Species: Kassina maculata
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Mar-1999
C;Accession: S10059
R;Yasuhara, T.; Nakajima, T.; Erspamer, G.F.; Erspamer, V.
Biomed. Res. 2, 613-617, 1981
A;Title: New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and hylambatin, i
A;Reference number: S07436
A;Accession: S10059
A;Molecule type: protein
A;Residues: 1-12 <YAS>
A;Experimental source: skin
A;Note: the source is designated as Hylambates maculatus
C;Keywords: amidated carboxyl end; neuropeptide; tachykinin
F;12/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 25.2%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 PRPDQ 12
|:|||
Db 3 PKPDQ 7

RESULT 2
JP0053
ribosomal protein L30 - Bacillus macerans (fragment)
C;Species: Bacillus macerans
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994
C;Accession: JP0053
R;Ochi, K.
submitted to JIPID, February 1994
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal
A;Reference number: JP0042
A;Accession: JP0053
A;Molecule type: protein
A;Residues: 1-20 <OCH>
C;Keywords: protein biosynthesis; ribosome

Query Match 25.2%; Score 27; DB 2; Length 20;
Best Local Similarity 35.7%; Pred. No. 2.5e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 2 MVKFLFPRPDQQHS 15
:|: | ||: | :
Db 7 LVRSLIGRPEXQRT 20

```
;       TELEFAX:  (206) 682-6031
;       TELEX:  3723836
;   INFORMATION FOR SEQ ID NO:  73:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-432-871C-73


  Query Match             25.2%;  Score 27;  DB 2;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    4;  Gaps    1;

Qy       5 FLFPRPDQQHSVAY 18
           |:| |     | :||
Db       1 FIFDR----HPIAY 10


RESULT  15
US-07-678-974D-8
; Sequence 8, Application US/07678974D
; Patent No. 5629146
;  GENERAL INFORMATION:
;     APPLICANT:  DILLNER, JOAKIM
;     APPLICANT:  DILLNER, LENA
;     TITLE OF INVENTION:  METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
;     TITLE OF INVENTION:
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BERMAN & AISENBERG
;       STREET:  1730 RHODE ISLAND AVENUE, N.W.,
;       CITY:  WASHINGTON
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20036-3186
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/678,974D
;       FILING DATE:  25-JUN-1991
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  AISENBERG, Irwin M.
;       REGISTRATION NUMBER:  19,007
;       REFERENCE/DOCKET NUMBER:  SG19171
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-293-1404
;       TELEFAX:  202-872-0493
;       TELEX:  440 069 AIS UI
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-07-678-974D-8


  Query Match             24.8%;  Score 26.5;  DB 1;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    5;  Gaps    1;

Qy       4 KFLFP-----RPDQQHSV 16
           || ||      || |  |
Db       2 KFGFPDTSFYNPDTQRLV 19


Search completed: December 21, 2000, 08:31:49
Job time: 372 sec
```

```
RESULT  12
US-08-221-580-7
; Sequence 7, Application US/08221580
; Patent No. 5519000
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Mervic, Miljenko
;    APPLICANT:  Weber, Robert W.
;    TITLE OF INVENTION:  Tumor Necrosis Factor Inhibitors
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5519000ris
;      STREET:  One Liberty Place 46th Floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  United States of America
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/221,580
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  CCOR-0140
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
;    MOLECULE TYPE:  peptide
US-08-221-580-7


  Query Match             25.2%;  Score 27;  DB 1;  Length 13:
  Best Local Similarity   55.6%;  Pred. No. 90;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 QQHSVAYTF 20
           | ||  :||
Db       5 QSHSWPFTF 13


RESULT  13
PCT-US95-04018-69
; Sequence 69, Application PC/TUS9504018
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Mervic, Miljenko
;    APPLICANT:  Weber, Robert W.
;    TITLE OF INVENTION:  Tumor Necrosis Factor Inhibitors
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place  46th Floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19403
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04018
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,580
;      FILING DATE:  01-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,583
;      FILING DATE:  01-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,581
;      FILING DATE:  01-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  CCOR-0232
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
;    MOLECULE TYPE:  peptide
PCT-US95-04018-69


  Query Match             25.2%;  Score 27;  DB 4;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 90;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 QQHSVAYTF 20
           | ||  :||
Db       5 QSHSWPFTF 13


RESULT  14
US-08-432-871C-73
; Sequence 73, Application US/08432871C
; Patent No. 5877010
;  GENERAL INFORMATION:
;    APPLICANT:  Loeb, Lawrence A.
;    APPLICANT:  Black, Margaret E.
;    TITLE OF INVENTION:  THYMIDINE KINASE MUTANTS
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/432,871C
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  240052.409C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,775
;      FILING DATE:  20-NOV-1997
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  7814/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-775-11


  Query Match             27.1%;  Score 29;  DB 3;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
            | |   |  ||
Db       3 PYPQGNHEAAY 13


RESULT  10
US-08-974-775-39
; Sequence 39, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,775
;      FILING DATE:  20-NOV-1997
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  7814/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-775-39


  Query Match             26.2%;  Score 28;  DB 3;  Length 14;
  Best Local Similarity   36.4%;  Pred. No. 66;
  Matches    4;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
            | |:  |  :|
Db       3 PYPEGNHEASY 13


RESULT  11
US-08-159-339A-558
; Sequence 558, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  558:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-558


  Query Match             25.2%;  Score 27;  DB 3;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 67;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 FPRPDQQH 14
            |  ||::|
Db       1 FTTPDKKH 8
```

```
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/606,639A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J.
;      REGISTRATION NUMBER:  36,677
;      REFERENCE/DOCKET NUMBER:  A-63005/PJS STAN-187
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 494-8700
;      TELEFAX:  (415) 494-8771
;      TELEX:  910 277299 FHT UR
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-606-639A-5


  Query Match             28.0%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   7
US-08-297-395-14
; Sequence 14, Application US/08297395A
; Patent No. 6039947
; GENERAL INFORMATION:
;  APPLICANT: Howard L. Weiner
;  APPLICANT: David A. Hafler
;  TITLE OF INVENTION: PEPTIDES DERIVED FROM IMMUNODOMINANT
;  TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN
;  FILE REFERENCE: 1010/05723US3
;  CURRENT APPLICATION NUMBER: US/08/297,395A
;  CURRENT FILING DATE: 1994-08-11
;  EARLIER APPLICATION NUMBER: 08/059,189
;  EARLIER FILING DATE: 1993-05-06
;  EARLIER APPLICATION NUMBER: 07/502,559
;  EARLIER FILING DATE: 1990-03-30
;  EARLIER APPLICATION NUMBER: PCT/US88/02139
;  EARLIER FILING DATE: 1988-06-24
;  EARLIER APPLICATION NUMBER: 07/065,734
;  EARLIER FILING DATE: 1987-06-24
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-297-395-14


  Query Match             28.0%;  Score 30;  DB 3;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   8
US-08-974-775-12
; Sequence 12, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,775
;      FILING DATE:  20-NOV-1997
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  7814/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-775-12


  Query Match             27.1%;  Score 29;  DB 3;  Length 13;
  Best Local Similarity   45.5%;  Pred. No. 40;
  Matches    5;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
           | |   |  ||
Db       3 PYPQGNHEAAY 13


RESULT   9
US-08-974-775-11
; Sequence 11, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COUNTRY:  USA
;       ZIP:  94105-1492
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/483,021
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/210,594
;       FILING DATE:  23-JUN-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/367,751
;       FILING DATE:  21-JUN-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/635,840
;       FILING DATE:  12-DEC-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/690,840
;       FILING DATE:  23-APR-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/793,938
;       FILING DATE:  19-NOV-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/978,946
;       FILING DATE:  18-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/208,072
;       FILING DATE:  03-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/397,533
;       FILING DATE:  02-MAR-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bastian, Kevin L.
;       REGISTRATION NUMBER:  34,774
;       REFERENCE/DOCKET NUMBER:  14058-16-1-4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 543-9600
;       TELEFAX:  (415) 543-5043
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /product= "Xaa = N-acetyl alanine"
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  1..14
;       OTHER INFORMATION:  /note= "rat myelin basic protein
;       OTHER INFORMATION:  peptide analog of MBP(1-14)4-A"
US-08-483-021-3


  Query Match             28.0%;  Score 30;  DB 1;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   5
US-08-468-540B-1
; Sequence 1, Application US/08468540B
; Patent No. 5858980
;  GENERAL INFORMATION:
;    APPLICANT:  Weiner, Howard
;    APPLICANT:  Hafler, David
;    APPLICANT:  Miller, Ariel
;    APPLICANT:  Al-Sabbagh, Ahmad
;    TITLE OF INVENTION:  SUPPRESSION OF T-CELL PROLIFERATION
;    TITLE OF INVENTION:  USING PEPTIDE FRAGMENTS OF MYELIN BASIC PROTEIN
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby P.C.
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,540B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jacobs, Seth H
;      REGISTRATION NUMBER:  32,140
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700'
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5858980e
US-08-468-540B-1


  Query Match             28.0%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   6
US-08-606-639A-5
; Sequence 5, Application US/08606639A
; Patent No. 5939400
;  GENERAL INFORMATION:
;    APPLICANT:  Steinman, Lawrence
;    APPLICANT:  Waisman, Ari
;    TITLE OF INVENTION:  DNA VACCINATION FOR INDUCTION OF
;    TITLE OF INVENTION:  SUPPRESSIVE T CELL RESPONSE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
```

```
  Query Match             31.8%;  Score 34;  DB 3;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 4;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 RPDQQHSVAY 18
          :||:|: |:|
Db      1 QPDKQYDVSY 10


RESULT   2
US-08-440-861-11
; Sequence 11, Application US/08440861
; Patent No. 5710126
;  GENERAL INFORMATION:
;    APPLICANT:  Griffith, Irwin J.
;    APPLICANT:  Kuo, Mei-Chang
;    APPLICANT:  Luqman, Mohammad
;    TITLE OF INVENTION:  T CELL EPITOPES OF RYEGRASS POLLEN
;    TITLE OF INVENTION:  ALLERGEN
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,861
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/106,016
;      FILING DATE:  31-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IPC-075 (IMI-040cp)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-440-861-11


  Query Match             29.0%;  Score 31;  DB 1;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 30;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      6 LFPRPDQQHSVAY 18
          | |: |   : |||
Db      3 LIPKLDTAYDVAY 15


RESULT   3
US-08-159-339A-1076
; Sequence 1076, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1076:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-1076


  Query Match             28.0%;  Score 30;  DB 3;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 20;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      9 RPDQQHSVAY 18
          || |:|   |
Db      1 RPSQRHGSKY 10


RESULT   4
US-08-483-021-3
; Sequence 3, Application US/08483021
; Patent No. 5734023
;  GENERAL INFORMATION:
;    APPLICANT:  Nag, Bishwajit
;    APPLICANT:  Clark, Brian R.
;    APPLICANT:  Sharma, Somesh
;    APPLICANT:  McConnell, Harden
;    TITLE OF INVENTION:  MHC Subunit Conjugates Useful in
;    TITLE OF INVENTION:  Ameliorating Deleterious Immune Responses
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:48 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-9
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	31.8	10	3	US-08-159-339A-935	Sequence 935, App
2	31	29.0	20	1	US-08-440-861-11	Sequence 11, Appl
3	30	28.0	10	3	US-08-159-339A-1076	Sequence 1076, Ap
4	30	28.0	14	1	US-08-483-021-3	Sequence 3, Appli
5	30	28.0	20	2	US-08-468-540B-1	Sequence 1, Appli
6	30	28.0	20	2	US-08-606-639A-5	Sequence 5, Appli
7	30	28.0	20	3	US-08-297-395-14	Sequence 14, Appl
8	29	27.1	13	3	US-08-974-775-12	Sequence 12, Appl
9	29	27.1	14	3	US-08-974-775-11	Sequence 11, Appl
10	28	26.2	14	3	US-08-974-775-39	Sequence 39, Appl
11	27	25.2	10	3	US-08-159-339A-558	Sequence 558, App
12	27	25.2	13	1	US-08-221-580-7	Sequence 7, Appli
13	27	25.2	13	4	PCT-US95-04018-69	Sequence 69, Appl
14	27	25.2	15	2	US-08-432-871C-73	Sequence 73, Appl
15	26.5	24.8	20	1	US-07-678-974D-8	Sequence 8, Appli
16	26.5	24.8	20	2	US-08-945-168-13	Sequence 13, Appl
17	26	24.3	8	3	US-08-159-339A-1049	Sequence 1049, Ap
18	26	24.3	8	3	US-08-925-002-49	Sequence 49, Appl
19	26	24.3	11	2	US-08-540-388-1	Sequence 1, Appli
20	26	24.3	12	1	US-07-778-233B-23	Sequence 23, Appl
21	26	24.3	12	1	US-07-963-321-23	Sequence 23, Appl
22	26	24.3	12	1	US-08-290-641-23	Sequence 23, Appl
23	26	24.3	12	1	US-08-548-540-23	Sequence 23, Appl
24	26	24.3	12	4	PCT-US96-09809-23	Sequence 23, Appl
25	26	24.3	13	1	US-08-469-005A-13	Sequence 13, Appl
26	26	24.3	13	1	US-08-188-426-7	Sequence 7, Appli
27	26	24.3	13	2	US-08-469-009-7	Sequence 7, Appli
28	26	24.3	13	2	US-08-469-007-7	Sequence 7, Appli
29	26	24.3	13	3	US-08-974-775-16	Sequence 16, Appl
30	26	24.3	13	3	US-08-462-351-8	Sequence 8, Appli
31	26	24.3	13	5	5194425-9	Patent No. 5194425
32	26	24.3	14	3	US-08-974-775-15	Sequence 15, Appl
33	26	24.3	18	2	US-08-951-871-10	Sequence 10, Appl
34	26	24.3	19	2	US-08-482-142-80	Sequence 80, Appl
35	26	24.3	19	2	US-08-613-235-6	Sequence 6, Appli
36	26	24.3	19	2	US-08-478-572-80	Sequence 80, Appl
37	25.5	23.8	20	2	US-08-107-676-21	Sequence 21, Appl
38	25	23.4	4	5	5425936-4	Patent No. 5425936
39	25	23.4	4	5	5433940-1	Patent No. 5433940
40	25	23.4	5	1	US-08-022-381A-12	Sequence 12, Appl
41	25	23.4	5	1	US-08-475-827A-12	Sequence 12, Appl
42	25	23.4	5	5	5196404-17	Patent No. 5196404
43	25	23.4	5	5	5433940-22	Patent No. 5433940
44	25	23.4	7	1	US-08-482-880-32	Sequence 32, Appl
45	25	23.4	7	2	US-08-273-274-32	Sequence 32, Appl

ALIGNMENTS

```
RESULT   1
US-08-159-339A-935
; Sequence 935, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  935:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-935
```

Search completed: December 21, 2000, 08:35:38
Job time: 601 sec

```
RT   enzymatic specificity and amino acid sequence deduced from mRNA.";
RL   Insect Mol. Biol. 3:201-211(1994).
SQ   SEQUENCE   18 AA;  1922 MW;  352EB0729B126B11 CRC64;


  Query Match             24.5%;  Score 25;  DB 5;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       5 QEVVGGFPSQ 14
           :  || || |
Db       6 EATVGQFPYQ 15


RESULT  12
Q06140
ID   Q06140      PRELIMINARY;      PRT;    19 AA.
AC   Q06140;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN P1.16 (PORIN)(FRAGMENT).
GN   PORA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93116032.
RA   Wedege E., Dalseg R., Caugant D.A., Poolman J.T., Froholm L.O.;
RT   "Expression of an inaccessible P1.7 subtype epitope on meningococcal
RT   class 1 proteins.";
RL   J. Med. Microbiol. 38:23-28(1993).
CC   -!- SUBUNIT: MONOMER.
DR   EMBL; S51901; AAB24741.1; -.
KW   Outer membrane; Antigen.
FT   NON_TER       1      1
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1773 MW;  AFAC572B037E0484 CRC64;


  Query Match             24.5%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+03;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 QEVVGGFPSQAQVT 18
           |   ||   | :||
Db       2 QAANGGASGQVKVT 15


RESULT  13
Q9R5I1
ID   Q9R5I1      PRELIMINARY;      PRT;    19 AA.
AC   Q9R5I1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   NITRILE HYDRATASE ALPHA SUBUNIT (FRAGMENT).
OS   Brevibacterium.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Brevibacteriaceae.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92407758.
RA   Duran R., Chion C.K., Bigey F., Arnaud A., Galzy P.;
RT   "The N-terminal amino acid sequences of Brevibacterium sp. R312
RT   nitrile hydratase.";
RL   J. Basic Microbiol. 32:13-19(1992).
SQ   SEQUENCE   19 AA;  1923 MW;  84726D1A1282FB63 CRC64;


  Query Match             24.5%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 PSQAQVT 18
           |:|| ||
Db      13 PAQAPVT 19


RESULT  14
Q9U8F7
ID   Q9U8F7      PRELIMINARY;      PRT;    18 AA.
AC   Q9U8F7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PAIRED DOMAIN TRANSCRIPTIONAL FACTOR EY (FRAGMENT).
GN   EYELESS.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu P.-X., Yoon A., Heaney S., Zhang X., Michelson A., Maas R.L.;
RT   "The regulation of Pax6 is conserved between mice and flies.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF089733; AAD54001.1; -.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1680 MW;  642A9C256071D610 CRC64;


  Query Match             23.5%;  Score 24;  DB 5;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 VGGFPS 13
           ||| |:
Db      12 VGGIPA 17


RESULT  15
Q62637
ID   Q62637      PRELIMINARY;      PRT;    19 AA.
AC   Q62637;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   XANTHINE DEHYDROGENASE (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE DAWLEY;
RX   MEDLINE; 94268906.
RA   Chow C.W., Clark M., Rinaldo J., Chalkley R.;
RT   "Identification of the rat xanthine dehydrogenase/oxidase promoter.";
RL   Nucleic Acids Res. 22:1846-1854(1994).
DR   EMBL; U08123; AAB60444.1; -.
FT   NON_TER       1      1
FT   VARIANT      19     19       H -> Q.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2128 MW;  A02841D686F50139 CRC64;


  Query Match             23.5%;  Score 24;  DB 11;  Length 19;
  Best Local Similarity   28.6%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVGGFPSQ 14
           :|: | |  |    :
Db       4 EELLQSVCAGLAEE 17
```

```
Db      "1 IVGGVKAQA 9


RESULT   7
O04817
ID   O04817      PRELIMINARY;      PRT;    19 AA.
AC   O04817;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   HYPOTHETICAL 2.3 KDA PROTEIN (FRAGMENT).
OS   Sporobolus stapfianus.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Sporobolus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Blomstedt C.K., Gianello R.D., Neale A.D., Hamill J.D., Gaff D.F.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y10781; CAA71753.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2270 MW;  4C644625813F244E CRC64;


  Query Match             25.5%;  Score 26;  DB 10;  Length 19;
  Best Local Similarity   27.3%;  Pred. No. 7.8e+02;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVGGF 11
           |::: ::  ||
Db       1 QKLYDDLTSGF 11


RESULT   8
Q9R4Z4
ID   Q9R4Z4      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AMBIENT-TEMPERATURE FIMBRIA; ATF.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Proteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94222573.
RA   Massad G., Bahrani F.K., Mobley H.L.;
RT   "Proteus mirabilis fimbriae: identification, isolation, and
RT   characterization of a new ambient-temperature fimbria.";
RL   Infect. Immun. 62:1989-1994(1994).
SQ   SEQUENCE   20 AA;  2017 MW;  B2FC481003883AE2 CRC64;


  Query Match             25.5%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQVTV 19
           | |:  :|||
Db       4 GTPAPTEVTV 13


RESULT   9
Q9UCA3
ID   Q9UCA3      PRELIMINARY;      PRT;    20 AA.
AC   Q9UCA3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PROSTASIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94308140.
RA   Yu J.X., Chao L., Chao J.;
RT   "Prostasin is a novel human serine proteinase from seminal fluid.
RT   Purification, tissue distribution, and localization in prostate
RT   gland.";
RL   J. Biol. Chem. 269:18843-18848(1994).
DR   INTERPRO; IPR001254; -.
DR   PFAM; PF00089; trypsin; 1.
SQ   SEQUENCE   20 AA;  2108 MW;  EDF142A1F9F880FE CRC64;


  Query Match             25.5%;  Score 26;  DB 4;  Length 20;
  Best Local Similarity   41.7%;  Pred. No. 8.3e+02;
  Matches    5;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVT 18
           | | :| |  :|
Db       8 VAGQWPWQVSIT 19


RESULT  10
Q9S8T0
ID   Q9S8T0      PRELIMINARY;      PRT;    20 AA.
AC   Q9S8T0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   JACALIN BETA-SUBUNIT.
OS   Artocarpus champeden.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Moraceae;
OC   Artocarpus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93152601.
RA   Ngoc L.D., Brillard M., Hoebeke J.;
RT   "The alpha- and beta-subunits of the jacalins are cleavage products
RT   from a 17-kDa precursor.";
RL   Biochim. Biophys. Acta 1156:219-222(1993).
SQ   SEQUENCE   20 AA;  2071 MW;  8D5AE975F4D0E212 CRC64;


  Query Match             25.0%;  Score 25.5;  DB 10;  Length 20;
  Best Local Similarity   56.2%;  Pred. No. 1e+03;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       3 IFQEVVGGFPSQAQVT 18
           | | |: | |  ||||
Db       6 ISQTVIVG-PWGAQVT 20


RESULT  11
Q9TWL4
ID   Q9TWL4      PRELIMINARY;      PRT;    18 AA.
AC   Q9TWL4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   25 KDA CHYMOTRYPSIN-LIKE ENZYME (FRAGMENT).
OS   Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Lucilia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95219141.
RA   Casu R.E., Pearson R.D., Jarmey J.M., Cadogan L.C., Riding G.A.,
RA   Tellam R.L.;
RT   "Excretory/secretory chymotrypsin from Lucilia cuprina: purification,
```

```
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: AUREIN 2.1 HAS ANTIMICROBIAL AND ANTICANCER ACTIVITY.
KW   Amidation; Antibiotic.
FT   PEPTIDE       3     16       AUREIN 2.1.1.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1616 MW;  1D9A5DADB4D240F9 CRC64;


  Query Match          28.4%;  Score 29;  DB 13;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           :| ::||| | |
Db       4 DIVKKVVGAFGS 15


RESULT   3
P82392
ID   P82392      PRELIMINARY;      PRT;    16 AA.
AC   P82392;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AUREIN 2.5.
OS   Litoria raniformis, and Litoria aurea (Australian frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: HAS ANTIMICROBIAL AND ANTICANCER ACTIVITY.
KW   Amidation; Antibiotic.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1650 MW;  1D9A5DADB4DAE2F9 CRC64;


  Query Match          28.4%;  Score 29;  DB 13;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           :| ::||| | |
Db       4 DIVKKVVGAFGS 15


RESULT   4
Q9TWK8
ID   Q9TWK8      PRELIMINARY;      PRT;    19 AA.
AC   Q9TWK8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   HAPTOGLOBIN-RELATED PROTEIN BETA SUBUNIT (FRAGMENT).
OS   Trypanosoma brucei.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95232503.
RA   Smith A.B., Esko J.D., Hajduk S.L.;
RT   "Killing of trypanosomes by the human haptoglobin-related protein.";
RL   Science 268:284-286(1995).
SQ   SEQUENCE   19 AA;  2055 MW;  65BD135667C94056 CRC64;


  Query Match          26.5%;  Score 27;  DB 5;  Length 19;
  Best Local Similarity   55.6%;  Pred. No. 5.3e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQV 17
           | || ||::
Db      10 GSFPWQAKM 18


RESULT   5
Q84274
ID   Q84274      PRELIMINARY;      PRT;    19 AA.
AC   Q84274;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   E6 PROTEIN (FRAGMENT).
OS   Human papillomavirus type 25.
OC   Viruses; dsDNA viruses, no RNA stage; Papovaviridae; Papillonavirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88089511.
RA   Krubke J., Kraus J., Delius H., Chow L., Broker T.R., Iftner T.,
RA   Pfister H.;
RT   "Genetic relationship among human papillomaviruses associated with
RT   benign and malignant tumours of patients with epidermodysplasia
RT   verruciformis.";
RL   J. Gen. Virol. 68:3091-3103(1987).
DR   EMBL; D00205; BAA00145.1; -.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1955 MW;  5C820487F69A5ED1 CRC64;


  Query Match          26.5%;  Score 27;  DB 12;  Length 19;
  Best Local Similarity   75.0%;  Pred. No. 5.3e+02;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQV 17
           | | ||||
Db      11 GPPEQAQV 18


RESULT   6
Q9TWV8
ID   Q9TWV8      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWV8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   TRYPSIN-LIKE PROTEASE (DER F III ALLERGEN HOMOLOG) (FRAGMENT).
OS   Dermatophagoides farinae (House-dust mite).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Acariformes; Sarcoptiformes; Astigmata; Analgoidea; Pyroglyphidae;
OC   Dermatophagoides.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93074427.
RA   Ando T., Homma R., Ino Y., Ito G., Miyahara A., Yamakawa H., Iwaki M.,
RA   Okumura Y., Suko M., Haida M.;
RT   "Is a trypsin-like protease of mites a Der f III allergen?";
RL   Jpn. J. Allergy 41:704-707(1992).
SQ   SEQUENCE   20 AA;  2087 MW;  961537685DB396A2 CRC64;


  Query Match          26.5%;  Score 27;  DB 5;  Length 20;
  Best Local Similarity   55.6%;  Pred. No. 5.7e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQA 15
           :|||  :||
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:38 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-8
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	29.4	20	6	Q9TR70	Q9tr70 ursus arcto
2	29	28.4	16	13	P82388	P82388 litoria ran
3	29	28.4	16	13	P82392	P82392 litoria ran
4	27	26.5	19	5	Q9TWK8	Q9twk8 trypanosoma
5	27	26.5	19	12	Q84274	Q84274 human papil
6	27	26.5	20	5	Q9TWV8	Q9twv8 dermatophag
7	26	25.5	19	10	O04817	O04817 sporobolus
8	26	25.5	20	2	Q9R4Z4	Q9r4z4 proteus mir
9	26	25.5	20	4	Q9UCA3	Q9uca3 homo sapien
10	25.5	25.0	20	10	Q9S8T0	Q9s8t0 artocarpus
11	25	24.5	18	5	Q9TWL4	Q9twl4 lucilia cup
12	25	24.5	19	2	Q06140	Q06140 neisseria m
13	25	24.5	19	2	Q9R5I1	Q9r5i1 brevibacter
14	24	23.5	18	5	Q9U8F7	Q9u8f7 drosophila
15	24	23.5	19	11	Q62637	Q62637 rattus norv
16	23	22.5	15	5	Q9TWF5	Q9twf5 artemia (br
17	23	22.5	16	13	P82389	P82389 litoria aur
18	23	22.5	16	13	P82390	P82390 litoria aur
19	23	22.5	17	4	Q9UCI3	Q9uci3 homo sapien
20	23	22.5	17	11	Q9QVC2	Q9qvc2 rattus sp.
21	23	22.5	18	6	Q9TT30	Q9tt30 saguinus la
22	23	22.5	18	10	Q9S892	Q9s892 glycine max
23	23	22.5	19	13	P82095	P82095 uperoleia m
24	23	22.5	20	4	Q9UD25	Q9ud25 homo sapien
25	22	21.6	14	11	Q10734	Q10734 mus musculu
26	22	21.6	15	12	Q88954	Q88954 vaccinia vi
27	22	21.6	16	2	Q9R514	Q9r514 porphyromon
28	22	21.6	17	6	Q9TRD5	Q9trd5 oryctolagus
29	22	21.6	18	4	Q14042	Q14042 homo sapien
30	22	21.6	19	2	Q9R5A9	Q9r5a9 arthrobacte
31	22	21.6	19	4	Q9UCG2	Q9ucg2 homo sapien
32	22	21.6	20	11	Q9QUZ7	Q9quz7 rattus sp.
33	22	21.6	20	13	Q9PRT8	Q9prt8 gadus morhu
34	22	21.6	20	13	Q9PRR3	Q9prr3 scyliorhinu
35	21.5	21.1	18	6	Q9TRB6	Q9trb6 sus scrofa
36	21.5	21.1	20	4	Q9UCM1	Q9ucm1 homo sapien
37	21	20.6	9	11	Q9QWT0	Q9qwt0 mus musculu
38	21	20.6	9	12	Q88953	Q88953 vaccinia vi
39	21	20.6	10	12	Q85462	Q85462 avian sarco
40	21	20.6	13	2	Q50038	Q50038 mycobacteri
41	21	20.6	15	11	Q9QUZ5	Q9quz5 mus sp. 38-
42	21	20.6	16	4	Q9UC48	Q9uc48 homo sapien
43	21	20.6	16	13	P82391	P82391 litoria aur
44	21	20.6	20	4	Q9UMU3	Q9umu3 homo sapien
45	21	20.6	20	6	Q9TRH6	Q9trh6 canis famil

ALIGNMENTS

RESULT 1
Q9TR70
ID Q9TR70 PRELIMINARY; PRT; 20 AA.
AC Q9TR70;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HAPTOGLOBIN BETA CHAIN (FRAGMENT).
OS Ursus arctos (Brown bear) (Grizzly bear).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Ursidae; Ursus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 95268704.
RA Mominoki K., Nakagawa-Tosa N., Morimatsu M., Syuto B., Saito M.;
RT "Haptoglobin in Carnivora: a unique molecular structure in bear, cat
RT and dog haptoglobins.";
RL Comp. Biochem. Physiol. Biochem. Mol. Biol. 110:785-789(1995).
SQ SEQUENCE 20 AA; 2020 MW; F605BCC82667DC3F CRC64;

Query Match 29.4%; Score 30; DB 6; Length 20;
Best Local Similarity 55.6%; Pred. No. 1.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy       9 GGFPSQAQV 17
           | || ||::
Db      10 GSFPXQAKM 18
```

RESULT 2
P82388
ID P82388 PRELIMINARY; PRT; 16 AA.
AC P82388;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE AUREIN 2.1/2.1.1.
OS Litoria raniformis, and Litoria aurea (Australian frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Litoria.

```
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   FERREDOXIN--NAD(+) REDUCTASE (EC 1.18.1.3) (FRAGMENT).
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92041607.
RA   Ramachandra M., Seetharam R., Emptage M.H., Sariaslani F.S.;
RT   "Purification and characterization of a soybean flour-inducible
RT   ferredoxin reductase of Streptomyces griseus.";
RL   J. Bacteriol. 173:7106-7112(1991).
CC   -!- FUNCTION: COUPLE ELECTRON TRANSFER FROM NADH TO CYTOCHROME
CC       P450(SOY) IN THE PRESENCE OF FERREDOXIN.
CC   -!- CATALYTIC ACTIVITY: REDUCED FERREDOXIN + NAD(+) = OXIDIZED
CC       FERREDOXIN + NADH.
CC   -!- COFACTOR: FAD; REQUIRES MAGNESIUM.
DR   HSSP; P09063; 1LVL.
KW   Oxidoreductase; Flavoprotein; NAD; FAD; Magnesium.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1485 MW;  27D11A7C37AC0510 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 9.5e+02;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 VVGGFP 12
           ::|| |
Db       5 IIGGGP 10


RESULT  14
FIXA_RHILE
ID   FIXA_RHILE     STANDARD;      PRT;    18 AA.
AC   P14313;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   FIXA PROTEIN (FRAGMENT).
GN   FIXA.
OS   Rhizobium leguminosarum.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Rhizobiaceae; Rhizobium.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 90066358.
RA   Hontelez J.G.J., Lankhorst R.K., Katinakis P., van den Bos R.C.,
RA   van Kammen A.;
RT   "Characterization and nucleotide sequence of a novel gene fixW
RT   upstream of the fixABC operon in Rhizobium leguminosarum.";
RL   Mol. Gen. Genet. 218:536-544(1989).
CC   -!- FUNCTION: MAY PLAY A ROLE IN A REDOX PROCESS INVOLVED IN NITROGEN
CC       FIXATION.
CC   -!- SUBUNIT: FIXA AND FIXB FORM A HETERODIMER (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X16521; CAA34527.1; -.
DR   PIR; PQ0022; PQ0022.
DR   INTERPRO; IPR000049; -.
DR   PROSITE; PS01065; ETF_BETA; PARTIAL.
KW   Electron transport; Nitrogen fixation.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2036 MW;  74973C8BA2087663 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
           |  ||: |
Db      11 PDSAQIRV 18


RESULT  15
TCBI_TRILO
ID   TCBI_TRILO     STANDARD;      PRT;    19 AA.
AC   P80070;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   TRICHOLONGIN BI AND BII.
OS   Trichoderma longibrachiatum.
OC   Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Trichoderma.
RN   [1]
RP   SEQUENCE, AND STRUCTURE BY NMR.
RX   MEDLINE; 92037622.
RA   Rebuffat S., Prigent Y., Auvin-Guette C., Bodo B.;
RT   "Tricholongins BI and BII, 19-residue peptaibols from Trichoderma
RT   longibrachiatum. Solution structure from two-dimensional NMR
RT   spectroscopy.";
RL   Eur. J. Biochem. 201:661-674(1991).
CC   -!- FUNCTION: TRICHOLONGINS ARE PEPTIDES WHICH DISPLAY ANTIFUNGAL
CC       AND ANTIBACTERIAL ACTIVITY. PROBABLY INTERACT WITH BIOLOGICAL
CC       MEMBRANES AND PROBABLY PRODUCES HOLES LEADING TO LEAKAGE.
CC   -!- PTM: AIB IS ALPHA-AMINO ISOBUTYRIC ACID.
CC   -!- PTM: IVA IS ISOVALINE.
KW   Antibiotic; Fungicide; Methylation; Acetylation; Hydroxylation;
KW   Membrane.
FT   MOD_RES       1      1       ACETYLATION.
FT   MOD_RES       1      1       METHYLATION (TO FORM AIB).
FT   MOD_RES       4      4       METHYLATION (TO FORM AIB).
FT   MOD_RES       5      5       METHYLATION (TO FORM AIB).
FT   MOD_RES       7      7       METHYLATION (TO FORM AIB).
FT   MOD_RES       8      8       METHYLATION (TO FORM AIB).
FT   MOD_RES       9      9       METHYLATION (TO FORM AIB).
FT   MOD_RES      12     12       METHYLATION (TO FORM AIB).
FT   MOD_RES      15     15       METHYLATION (TO FORM AIB).
FT   MOD_RES      16     16       METHYLATION (TO FORM AIB; IN BI).
FT   MOD_RES      16     16       ETHYLATION (TO FORM IVA; IN BII).
FT   MOD_RES      19     19       HYDROXYLATION.
SQ   SEQUENCE   19 AA;  1756 MW;  B908C2DFF883D238 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQVTV 19
           || :||  ::
Db       2 GFAAQAAASL 11


Search completed: December 21, 2000, 08:32:58
Job time: 441 sec
```

```
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "New antibiotic uperin peptides from the dorsal glands of the
RT   australian toadlet Uperoleia mjobergii.";
RL   Aust. J. Chem. 49:1325-1331(1996).
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1948; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   19 AA;  1949 MW;  24E4F83A6BA35F21 CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGG 10
           :| :::|||
Db       4 DIAKKLVGG 12


RESULT  10
TRYL_STREX
ID   TRYL_STREX     STANDARD;      PRT;     20 AA.
AC   P80420;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   TRYPSIN-LIKE PROTEASE (EC 3.4.21.-) (FRAGMENT).
OS   Streptomyces exfoliatus (Streptomyces hydrogenans).
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=SMF13;
RX   MEDLINE; 95291424.
RA   Kim I.S., Lee K.J.;
RT   "Physiological roles of leupeptin and extracellular proteases in
RT   mycelium development of Streptomyces exfoliatus SMF13.";
RL   Microbiology 141:1017-1025(1995).
CC   -!- FUNCTION: INVOLVED IN MYCELIUM DIFFERENTIATION.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A; ALSO KNOWN AS THE
CC       ALPHA-LYTIC PROTEASE FAMILY.
DR   HSSP; P00775; 1SGT.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2129 MW;  4568F56D0E7393AF CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   36.4%;  Pred. No. 8.2e+02;
  Matches    4;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           | || |  :::
Db      10 GNFPFQQXLSM 20


RESULT  11
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;     9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83160932.
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   INTERPRO; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN        1      >9       PENTAXIN.
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 GFPSQAQV 17
           ||| :: :
Db       1 GFPGKSLI 8


RESULT  12
UN46_CLOPA
ID   UN46_CLOPA     STANDARD;      PRT;     14 AA.
AC   P81362;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   UNKNOWN PROTEIN CP 46 FROM 2D-PAGE (FRAGMENT).
OS   Clostridium pasteurianum.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=W5;
RX   MEDLINE; 98291870.
RA   Flengsrud R., Skjeldal L.;
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 5.4, ITS MW IS: 38.2 KDA.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1550 MW;  198078F4C0367170 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   42.9%;  Pred. No. 8.3e+02;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 IFQEVVG 9
           || :::|
Db       2 IFNDLIG 8


RESULT  13
FENR_STRGR
ID   FENR_STRGR     STANDARD;      PRT;     16 AA.
AC   P24134;
DT   01-MAR-1992 (Rel. 21, Created)
```

```
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGG 10
           :: ::||||
Db       4 DLAKKVVGG 12


RESULT   6
UP23_UPEIN
ID   UP23_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82029;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.3.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST B.CEREUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1974; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1975 MW;  2F34EF077BA35B70 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 5.2e+02;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGG 10
           :: ::||||
Db       4 DLAKKVVGG 12


RESULT   7
COG1_CHIOP
ID   COG1_CHIOP     STANDARD;      PRT;    20 AA.
AC   P34153;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   COLLAGENOLYTIC PROTEASE 25 KDA II/III (EC 3.4.21.32) (FRAGMENT).
OS   Chionoecetes opilio (Crab-beetle).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Majoidea; Majidae; Chionoecetes.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEPATOPANCREAS;
RX   MEDLINE; 92120073.
RA   Klimova O.A., Vedishcheva Y.V., Strongin A.Y.;
RT   "Isolation and characteristics of collagenolytic enzymes from the
RT   hepatopancreas of the crab Chionoecetes opilio.";
RL   Dokl. Akad. Nauk SSSR 317:482-484(1991).
CC   -!- FUNCTION: THIS ENZYME IS A SERINE PROTEASE CAPABLE OF DEGRADING
CC       THE NATIVE TRIPLE HELIX OF COLLAGEN.
CC   -!- CATALYTIC ACTIVITY: BROAD SPECIFICITY; DEGRADES NATIVE COLLAGEN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   HSSP; P00771; 1AZZ.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Collagen degradation.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2204 MW;  CE0D7B996E7281A7 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   35.7%;  Pred. No. 5.5e+02;
  Matches    5;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVTVH 20
           :|||  :     ||
Db       1 IVGGQEATPHTWVH 14


RESULT   8
CPAX_BOVIN
ID   CPAX_BOVIN     STANDARD;      PRT;    18 AA.
AC   P22779;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CYTOCHROME P450 2A (OLF2) (OLFACTIVE) (P52) (EC 1.14.14.1) (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 91027757.
RA   Lazard D., Tal N., Rubinstein M., Khen M., Lancet D., Zupko K.;
RT   "Identification and biochemical analysis of novel olfactory-specific
RT   cytochrome P-450IIA and UDP-glucuronosyl transferase.";
RL   Biochemistry 29:7433-7440(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A35704; A35704.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum; Olfaction.
FT   NON_TER       1      1
FT   VARIANT       6      6       G -> D.
FT   VARIANT      11     11       A -> E.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2058 MW;  F80746F76CCD77FF CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 7.3e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVG 9
           |: |:|: |
Db       9 QQAFKELQG 17


RESULT   9
UP27_UPEMJ
ID   UP27_UPEMJ     STANDARD;      PRT;    19 AA.
AC   P82039;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.7.
OS   Uperoleia mjobergii (Australian toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
RESULT   2
COCO_LIMPO
ID   COCO_LIMPO     STANDARD;      PRT;    14 AA.
AC   P35586;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   COCOONASE (EC 3.4.21.-) (FRAGMENT).
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 78037243.
RA   Law J.H., Dunn P.E., Kramer K.J.;
RT   "Insect proteases and peptidases.";
RL   Adv. Enzymol. Relat. Areas Mol. Biol. 45:389-425(1977).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   HSSP; P00760; 4TPI.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1452 MW;  1615FB1D73747570 CRC64;


  Query Match             24.5%;  Score 25;  DB 1;  Length 14;
  Best Local Similarity   80.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT   3
TRYP_FELCA
ID   TRYP_FELCA     STANDARD;      PRT;    16 AA.
AC   P81071;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT).
OS   Felis silvestris catus (Cat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PANCREAS;
RX   MEDLINE; 97235546.
RA   Steiner J.M., Medinger T.L., Williams D.A.;
RT   "Purification and partial characterization of feline trypsin.";
RL   Comp. Biochem. Physiol. 116B:87-93(1997).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen.
FT   PROPEP        1      8       ACTIVATION PEPTIDE.
FT   CHAIN         9    >16       TRYPSIN.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1825 MW;  A6D751BB58760A86 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 16;
  Best Local Similarity   33.3%;  Pred. No. 4.4e+02;
  Matches    3;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 EVVGGFPSQ 14
           ::|||: ::
Db       8 KIVGGYTNR 16


RESULT   4
ALL7_OLEEU
ID   ALL7_OLEEU     STANDARD;      PRT;    19 AA.
AC   P81430;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   POLLEN ALLERGEN OLE E 7 (OLE E VII) (FRAGMENT).
OS   Olea europaea (Common olive).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC   Lamiales; Oleaceae; Olea.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=POLLEN;
RA   Tejera M.L., Villalba M., Rodriguez R.;
RT   "Isolation and characterization of Ole e 7, a new allergen from olive
RT   tree pollen.";
RL   Submitted (JUL-1998) to the SWISS-PROT data bank.
KW   Allergen.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1986 MW;  999A5008C414A67E CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity   71.4%;  Pred. No. 5.2e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 PSQAQVT 18
           |||  ||
Db       2 PSQGTVT 8


RESULT   5
UP22_UPEIN
ID   UP22_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82028;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.2.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A WEAK ANTIBACTERIAL ACTIVITY AGAINST B.CEREUS,
CC       E.COLI, L.MESENTERIODES, L.INNOCUA, M.LUTEUS, P.HAEMOLYTICA,
CC       S.AUREUS AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1927 MW;  3283EF077BA35B70 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 5.2e+02;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:57 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-8
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	25.5	15	1	GLN2_PINPS	P81107 pinus pinas
2	25	24.5	14	1	COCO_LIMPO	P35586 limulus pol
3	24	23.5	16	1	TRYP_FELCA	P81071 felis silve
4	24	23.5	19	1	ALL7_OLEEU	P81430 olea europa
5	24	23.5	19	1	UP22_UPEIN	P82028 uperoleia i
6	24	23.5	19	1	UP23_UPEIN	P82029 uperoleia i
7	24	23.5	20	1	COG1_CHIOP	P34153 chionoecete
8	23	22.5	18	1	CPAX_BOVIN	P22779 bos taurus
9	23	22.5	19	1	UP27_UPEMJ	P82039 uperoleia m
10	23	22.5	20	1	TRYL_STREX	P80420 streptomyce
11	22	21.6	9	1	SAMP_MUSCA	P19095 mustelus ca
12	22	21.6	14	1	UN46_CLOPA	P81362 clostridium
13	22	21.6	16	1	FENR_STRGR	P24134 streptomyce
14	22	21.6	18	1	FIXA_RHILE	P14313 rhizobium l
15	22	21.6	19	1	TCBI_TRILO	P80070 trichoderma
16	22	21.6	19	1	UP21_UPEIN	P82027 uperoleia i
17	21.5	21.1	20	1	COGC_PARCM	P20734 paralithode
18	21	20.6	17	1	UP37_UPEMJ	P82044 uperoleia m
19	21	20.6	18	1	AROF_STRRM	P80575 streptomyce
20	21	20.6	19	1	OXLA_OPHHA	P81383 ophiophagus
21	21	20.6	20	1	ATP4_SPIOL	P80085 spinacia ol
22	21	20.6	20	1	ELAS_GADMO	P32197 gadus morhu
23	21	20.6	20	1	TL18_SPIOL	P82536 spinacia ol
24	20	19.6	13	1	BPP1_BOTJA	P01020 bothrops ja
25	20	19.6	13	1	UN12_CLOPA	P81353 clostridium
26	20	19.6	15	1	CYSK_CLOPA	P81340 clostridium
27	20	19.6	15	1	MILT_ONCKE	P81037 oncorhynchu
28	20	19.6	15	1	UC25_MAIZE	P80631 zea mays (m
29	20	19.6	16	1	CT12_LITCI	P81840 litoria cit
30	20	19.6	16	1	LPK1_LOCMI	P20404 locusta mig
31	20	19.6	17	1	EFG_THEAQ	Q01697 thermus aqu
32	20	19.6	19	1	UP24_UPEIN	P82030 uperoleia i
33	20	19.6	20	1	ANCR_AGKBI	P33588 agkistrodon
34	19	18.6	10	1	FARP_MANSE	P18523 manduca sex
35	19	18.6	13	1	CRBL_VESTR	P17231 vespa tropi
36	19	18.6	15	1	ATP2_PINPS	P81663 pinus pinas
37	19	18.6	15	1	COXJ_THUOB	P80979 thunnus obe
38	19	18.6	15	1	ESTJ_MANSE	P19985 manduca sex
39	19	18.6	15	1	FIBA_ANAPL	P12801 anas platyr
40	19	18.6	16	1	CT11_LITCI	P81835 litoria cit
41	19	18.6	16	1	CT13_LITCI	P81846 litoria cit
42	19	18.6	19	1	HBB2_UROHA	P18992 uromastyx h
43	19	18.6	20	1	ALAT_PIG	P13191 sus scrofa
44	19	18.6	20	1	COG1_PARCM	P20731 paralithode
45	19	18.6	20	1	COG2_CHIOP	P34154 chionoecete

ALIGNMENTS

RESULT 1
GLN2_PINPS
ID GLN2_PINPS STANDARD; PRT; 15 AA.
AC P81107;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROBABLE GLUTAMINE SYNTHETASE LEAF ISOZYME (EC 6.3.1.2) (GLUTAMATE--
DE AMMONIA LIGASE) (S2205/S2287) (N47/N48) (FRAGMENT).
OS Pinus pinaster (Maritime pine).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Coniferopsida; Coniferales; Pinaceae; Pinus.
RN [1]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RA Plomion C., Costa P., Bahrman N., Frigerio J.-M.;
RT "Genetic analysis of needle proteins in maritime pine. 1. Mapping
RT dominant and codominant protein markers assayed on diploid tissue, in
RT a haploid-based genetic map.";
RL Silvae Genetica 46:161-165(1997).
RN [2]
RP SEQUENCE.
RC TISSUE=NEEDLE;
RX MEDLINE; 99274088.
RA Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA Frigerio J.-M., Plomion C.;
RT "Separation and characterization of needle and xylem maritime pine
RT proteins.";
RL Electrophoresis 20:1098-1108(1999).
CC -!- FUNCTION: THE LIGHT-MODULATED CHLOROPLAST ENZYME, ENCODED BY A
CC NUCLEAR GENE AND EXPRESSED PRIMARILY IN LEAVES, IS RESPONSIBLE FOR
CC THE REASSIMILATION OF THE AMMONIA GENERATED BY PHOTORESPIRATION
CC (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: ATP + L-GLUTAMATE + NH(3) = ADP + GLUTAMINE +
CC ORTHOPHOSPHATE.
CC -!- SUBUNIT: HOMOOCTAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC 5.7, ITS MW IS: 42 KDA.
CC -!- SIMILARITY: BELONGS TO THE GLUTAMINE SYNTHETASE FAMILY.
DR INTERPRO; IPR001691; -.
DR PROSITE; PS00180; GLNA_1; PARTIAL.
DR PROSITE; PS00181; GLNA_ATP; PARTIAL.
KW Ligase; Chloroplast.
FT NON_TER 1 1
FT NON_TER 15 15
SQ SEQUENCE 15 AA; 1614 MW; 24A2420BEFD60D27 CRC64;

Query Match 25.5%; Score 26; DB 1; Length 15;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       8 VGGFP 12
           |||:|
Db      10 VGGYP 14
```

```
Db      3 VSGGFAS 9


RESULT  14
PA0026
protein QA300027 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997
C;Accession: PA0026
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensiona
A;Reference number: PA0001
A;Accession: PA0026
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: leaf


  Query Match            22.5%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 1.1e+03;
  Matches    5;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      8 VGGFPSQAQVT 18
          | | |   |:|
Db      4 VYGXPFAXQIT 14


RESULT  15
A35704
cytochrome P450 olf2 - bovine (fragment)
N;Contains: oxidoreductase (EC 1.-.-.-)
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 12-Mar-1999
C;Accession: A35704
R;Lazard, D.; Tal, N.; Rubinstein, M.; Khen, M.; Lancet, D.; Zupko, K.
Biochemistry 29, 7433-7440, 1990
A;Title: Identification and biochemical analysis of novel olfactory-specific cytochrome
A;Reference number: A35704; MUID:91027757
A;Accession: A35704
A;Molecule type: protein
A;Residues: 1-18 <LAZ>
C;Genetics:
A;Gene: CYP2A
C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C;Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protei


  Query Match            22.5%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 QEIFQEVVG 9
          |: |:|: |
Db      9 QQAFKELQG 17


Search completed: December 21, 2000, 08:30:06
Job time: 270 sec
```

```
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             24.5%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 VGGFPS 13
           |||| |
Db       5 VGGFAS 10


RESULT   9
A20190
hypodermin B - early cattle grub (fragment)
C;Species: Hypoderma lineatum (early cattle grub)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 18-Jun-1993
C;Accession: A20190
R;Lecroisey, A.; Tong, N.T.; Keil, B.
Eur. J. Biochem. 134, 261-267, 1983
A;Title: Hypodermin B, a trypsin-related enzyme from the insect Hypoderma lineatum.
A;Reference number: A20190; MUID:83261874
A;Accession: A20190
A;Molecule type: protein
A;Residues: 1-16 <LEC>


  Query Match             24.5%;  Score 25;  DB 2;  Length 16;
  Best Local Similarity   80.0%;  Pred. No. 5.5e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT  10
A45806
T-cell receptor beta chain C region type 1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 23-Jul-1999
C;Accession: A45806
R;Dent, A.L.; Fink, P.J.; Hedrick, S.M.
J. Immunol. 143, 322-328, 1989
A;Title: Characterization of an alternative exon of the murine T cell receptor beta-chai
A;Reference number: A45806; MUID:89278666
A;Accession: A45806
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-20 <DEN>
A;Cross-references: GB:M27225; NID:g339373; PIDN:AAA61099.1; PID:g553781
C;Superfamily: immunoglobulin C region; immunoglobulin homology
C;Keywords: T-cell receptor


  Query Match             24.0%;  Score 24.5;  DB 2;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    5;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       2 EIFQEVVGGF-PSQAQV 17
           ::|   |  | ||:|::
Db       4 KVFPPEVAVFEPSEAEI 20


RESULT  11
I52698
hypothetical THRA1/BTR mutant fusion protein, cell line BT474 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 20-Apr-2000
C;Accession: I52698
R;Futreal, P.A.; Cochran, C.; Marks, J.R.; Iglehart, J.D.; Zimmerman, W.; Barrett, J.C.;
Cancer Res. 54, 1791-1794, 1994
A;Title: Mutation analysis of the THRA1 gene in breast cancer: deletion/fusion of the
A;Reference number: I52698; MUID:94185019
A;Accession: I52698
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-15 <FUT>
A;Cross-references: GB:S71020; NID:g546111; PIDN:AAB30341.1; PID:g546112
C;Comment: This sequence is the chimeric product of a deletion or translocation mutat
C;Genetics:
A;Gene: THRA1/BTR
A;Map position: 17q11.2
C;Keywords: fusion protein


  Query Match             23.5%;  Score 24;  DB 4;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 7.5e+02;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 FPSQAQVTVH 20
           |  | :| ||
Db       3 FSEQFRVQVH 12


RESULT  12
PC2241
heat shock protein 42A - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 24-Feb-1995 #sequence_revision 24-Feb-1995 #text_change 17-Mar-1999
C;Accession: PC2241
R;Hatayama, T.; Yasuda, K.; Nishiyama, E.
Biochem. Biophys. Res. Commun. 204, 357-365, 1994
A;Title: Characterization of high-molecular-mass heat shock proteins and 42oC-specifi
A;Reference number: PC2238; MUID:95032120
A;Accession: PC2241
A;Molecule type: protein
A;Residues: 1-17 <HAT>
C;Keywords: heat shock; stress-induced protein


  Query Match             23.5%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   80.0%;  Pred. No. 8.7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 EVVGG 10
           |:|||
Db       8 EIVGG 12


RESULT  13
PH1448
T-cell receptor alpha chain (clone A24/PEG2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1448
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.;
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompa
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1448
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             22.5%;  Score 23;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 1e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 VVGGFPS 13
           | ||| |
```

```
RESULT   3
S56122
type I DNA methyltransferase M.EcoR124 I chain HsdM - Escherichia coli (fragments)
C;Species: Escherichia coli
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 03-Nov-1995
C;Accession: S56122
R;Webb, M.; Taylor, I.A.; Firman, K.; Kneale, G.G.
J. Mol. Biol. 250, 181-190, 1995
A;Title: Probing the domain structure of the type IC DNA methyltransferase M.EcoR124I by
A;Reference number: S56121; MUID:95333175
A;Accession: S56122
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-12 <WEB>


  Query Match             25.5%;  Score 26;  DB 2;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 GFPSQA 15
           |:||:|
Db       7 GYPSEA 12


RESULT   4
A61392
brain-associated small cell lung cancer antigen - human (fragment)
N;Alternate names: BASCA
C;Species: Homo sapiens (man)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C;Accession: A61392
R;Umezawa, Y.; Kuge, S.; Kikyo, N.; Shirai, T.; Watanabe, J.; Fujiwara, M.; Okabe, T.
Jpn. J. Clin. Oncol. 21, 251-255, 1991
A;Title: Identity of brain-associated small cell lung cancer antigen and the CD56 (NKH-1
A;Reference number: A61392; MUID:92046737
A;Accession: A61392
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <UME>


  Query Match             25.5%;  Score 26;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
           ||| :::|
Db       7 PSQGEISV 14


RESULT   5
C34917
catechol 1,2-dioxygenase (EC 1.13.11.1) beta chain - Pseudomonas sp. (fragment)
N;Alternate names: pyrocatechase beta chain
C;Species: Pseudomonas sp.
C;Date: 20-Jul-1990 #sequence_revision 20-Jul-1990 #text_change 10-Feb-1995
C;Accession: C34917
R;Nakai, C.; Horiike, K.; Kuramitsu, S.; Kagamiyama, H.; Nozaki, M.
J. Biol. Chem. 265, 660-665, 1990
A;Title: Three isozymes of catechol 1,2-dioxygenase (pyrocatechase), alphaalpha, alphabe
A;Reference number: A34917; MUID:90110118
A;Accession: C34917
A;Molecule type: protein
A;Residues: 1-20 <NAK>
C;Keywords: heterodimer; homodimer; oxidoreductase


  Query Match             25.5%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 4.8e+02;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVG 9
             |: |:|  |
Db      11 QQFFEEAAG 19


RESULT   6
S29636
jacalin beta-I chain - Artocarpus champeden (fragment)
C;Species: Artocarpus champeden
C;Date: 19-Mar-1997 #sequence_revision 17-Jul-1998 #text_change 17-Jul-1998
C;Accession: S29636
R;Ngoc, L.D.; Brillard, M.; Hoebeke, J.
Biochim. Biophys. Acta 1156, 219-222, 1993
A;Title: The alpha- and beta-subunits of the jacalins are cleavage products from a 17
A;Reference number: S29635; MUID:93152601
A;Accession: S29636
A;Molecule type: protein
A;Residues: 1-20 <NGO>
A;Experimental source: seed
C;Complex: heterotetramer; two alpha and two beta chains
C;Function:
A;Description: seed storage protein
A;Note: lectin for D-galactosyl-beta-1->3-N-acetylgalactosamine, a tumor-associated T
C;Keywords: heterotetramer; lectin; seed; storage protein


  Query Match             25.0%;  Score 25.5;  DB 2;  Length 20;
  Best Local Similarity   56.2%;  Pred. No. 5.9e+02;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       3 IFQEVVGGFPSQAQVT 18
           | |  |: | |  ||||
Db       6 ISQTVIVG-PWGAQVT 20


RESULT   7
B61168
cocoonase (EC 3.4.21.-) - Chinese oak silkmoth (fragment)
C;Species: Antheraea pernyi (Chinese oak silkmoth)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C;Accession: B61168
R;Kramer, K.J.; Felsted, R.L.; Law, J.H.
J. Biol. Chem. 248, 3021-3028, 1973
A;Title: Cocoonase. V. Structural studies on an insect serine protease.
A;Reference number: A61168; MUID:73166540
A;Accession: B61168
A;Molecule type: protein
A;Residues: 1-5 <KRA>
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-5/Product: cocoonase (fragment) #status experimental <MAT>


  Query Match             24.5%;  Score 25;  DB 2;  Length 5;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+05;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 VVGGF 11
           :||||
Db       1 IVGGF 5


RESULT   8
PH1436
T-cell receptor alpha chain (clone A24/PEF5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1436
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.;
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompa
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1436
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:06 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-8
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	28.4	16	2	A60551	leukocyte elastase
2	27	26.5	17	2	A61557	major merozoite su
3	26	25.5	12	2	S56122	type I DNA methylt
4	26	25.5	18	2	A61392	brain-associated s
5	26	25.5	20	2	C34917	catechol 1,2-dioxy
6	25.5	25.0	20	2	S29636	jacalin beta-I cha
7	25	24.5	5	2	B61168	cocoonase (EC 3.4.
8	25	24.5	15	2	PH1436	T-cell receptor al
9	25	24.5	16	2	A20190	hypodermin B - ear
10	24.5	24.0	20	2	A45806	T-cell receptor be
11	24	23.5	15	4	I52698	hypothetical THRA1
12	24	23.5	17	2	PC2241	heat shock protein
13	23	22.5	14	2	PH1448	T-cell receptor al
14	23	22.5	15	2	PA0026	protein QA300027 -
15	23	22.5	18	2	A35704	cytochrome P450 ol
16	23	22.5	20	2	S50022	trypsin-like prote
17	23	22.5	20	2	PU0033	aldose 1-epimerase
18	22.5	22.1	20	2	A56900	chymotrypsin I (EC
19	22	21.6	9	2	B20569	serum amyloid P-co
20	22	21.6	12	2	PH1183	T-cell receptor al
21	22	21.6	14	2	PH0747	T-cell receptor be
22	22	21.6	15	2	S26527	T-cell receptor al
23	22	21.6	16	2	C53113	Lys-gingipain form
24	22	21.6	16	2	A44897	ferredoxin--NADP+
25	22	21.6	17	2	B61334	trypsin (EC 3.4.21
26	22	21.6	18	2	PQ0022	fixA protein - Rhi
27	22	21.6	19	2	S29167	quinaldine oxidore
28	22	21.6	19	2	S74087	antibacterial prot
29	22	21.6	20	2	S50023	trypsin-like prote
30	22	21.6	20	2	S71601	recombination prot
31	21.5	21.1	20	2	S29108	glutathione transf
32	21.5	21.1	20	2	A61327	trypsin (EC 3.4.21
33	21.5	21.1	20	2	D34817	collagenolytic pro
34	21	20.6	12	2	PH1189	T-cell receptor al
35	21	20.6	12	2	PH1180	T-cell receptor al
36	21	20.6	12	2	PH1188	T-cell receptor al
37	21	20.6	12	2	PH1172	T-cell receptor al
38	21	20.6	12	2	PH1175	T-cell receptor al
39	21	20.6	13	2	A32734	enkephalin precurs
40	21	20.6	14	2	PA0015	seed storage prote
41	21	20.6	14	2	PS0278	ribulose-bisphosph
42	21	20.6	14	2	PH1450	T-cell receptor al
43	21	20.6	15	2	S26516	T-cell receptor al
44	21	20.6	15	2	S26524	T-cell receptor al
45	21	20.6	15	2	S26528	T-cell receptor al

ALIGNMENTS

RESULT 1
A60551
leukocyte elastase (EC 3.4.21.37) - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 17-Mar-1999
C;Accession: A60551
R;Axelsson, L.; Bergenfeldt, M.; Bjoerk, P.; Olsson, R.; Ohlsson, K.
Scand. J. Clin. Lab. Invest. 50, 35-42, 1990
A;Title: Release of immunoreactive canine leukocyte elastase normally and in endotoxi
A;Reference number: A60551; MUID:90193608
A;Accession: A60551
A;Molecule type: protein
A;Residues: 1-16 <AXE>
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; leukocyte; lysosome; serine proteinase

Query Match 28.4%; Score 29; DB 2; Length 16;
Best Local Similarity 62.5%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy       7 VVGGFPSQ 14
            :||| |:|
Db       1 IVGGRPAQ 8
```

RESULT 2
A61557
major merozoite surface antigen - malaria parasite (Plasmodium falciparum) (strain Fc
C;Species: Plasmodium falciparum
C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jun-2000
C;Accession: A61557
R;Heidrich, H.G.
Biol. Cell 64, 205-214, 1988
A;Title: Isolation and functional characterization of Plasmodium falciparum merozoite
A;Reference number: A61557; MUID:89150734
A;Accession: A61557
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-17 <HEI>
C;Keywords: surface antigen

Query Match 26.5%; Score 27; DB 2; Length 17;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy       2 EIFQEVVGGFPS 13
            | :||:|   ||
Db       4 ESYQELVAVTPS 15
```

```
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive, 6300 Sear Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/286,889
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/173,497
;       FILING DATE:  23-DEC-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Williams Jr., Joseph A.
;       REGISTRATION NUMBER:  P38,659
;       REFERENCE/DOCKET NUMBER:  27866/32168
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-474-6300
;       TELEFAX:  312-474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:. 11 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-286-889-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
           :|||   ||
Db       1 VFQEXGAGF 9


RESULT  15
US-08-485-618-15
; Sequence 15, Application US/08485618
; Patent No. 5728533
;  GENERAL INFORMATION:
;     APPLICANT:  Gallatin, W. Michael
;     APPLICANT:  Van der Vieren, Monica
;     TITLE OF INVENTION:  No. 5728533el Human  2 Integrin Alpha Subunit
;     NUMBER OF SEQUENCES:  103
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  233 South Wacker Drive, 6300 Sear Tower
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,618
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/173,497
;       FILING DATE:  23-DEC-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/286,889
;       FILING DATE:  5-AUG-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/362,652
;       FILING DATE:  21-DEC-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Williams Jr., Joseph A.
;       REGISTRATION NUMBER:  38,659
;       REFERENCE/DOCKET NUMBER:  27866/32797
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-474-6300
;       TELEFAX:  312-474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  15:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  11 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-485-618-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
           :|||   ||
Db       1 VFQEXGAGF 9


Search completed: December 21, 2000, 08:31:48
Job time: 371 sec
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-666-473-114


  Query Match             27.5%;  Score 28;  DB 2;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 PSQAQVTV 19
           |  |||||
Db       1 PHMAQVTV 8


RESULT  12
US-08-669-284B-26
; Sequence 26, Application US/08669284B
; Patent No. 5939534
;  GENERAL INFORMATION:
;    APPLICANT:  Inoue, Makoto
;    APPLICANT:  Kikuchi, Kaoru
;    APPLICANT:  Ishige, Yoko
;    APPLICANT:  Ito, Akira
;    APPLICANT:  Kimura, Toru
;    APPLICANT:  Nakayama, Chikao
;    APPLICANT:  No. 5939534uchi, Hiroshi
;    TITLE OF INVENTION:  NOVEL HUMAN CILIARY NEUROTROPHIC FACTORS
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/669,284B
;      FILING DATE:  28-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP94/02269
;      FILING DATE:  27-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 06-268281
;      FILING DATE:  05-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 06-201504
;      FILING DATE:  02-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 05-350934
;      FILING DATE:  29-DEC-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Nakamura, Dean H.
;      REGISTRATION NUMBER:  33,981
;      REFERENCE/DOCKET NUMBER:  Q-42041
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  8 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-669-284B-26


  Query Match             26.5%;  Score 27;  DB 2;  Length 8;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 FQEVVGG 10
            ||: |||
Db       2 FQQQVGG 8


RESULT  13
US-08-173-497-15
; Sequence 15, Application US/08173497
; Patent No. 5437958
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van Der Vieren, Monica
;    TITLE OF INVENTION:  No. 5437958el Human  2 Integrin Alpha
;    TITLE OF INVENTION:  Subunit
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 S. Wacker Drive, 6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/173,497
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5437958and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  27866/31363
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-173-497-15


  Query Match             26.5%;  Score 27;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGGF 11
            :|||   ||
Db       1 VFQEXGAGF 9


RESULT  14
US-08-286-889-15
; Sequence 15, Application US/08286889
; Patent No. 5470953
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Mich
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5470953el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
```

```
;    TITLE OF INVENTION:  METHODS FOR THE TREATMENT OF
;    TITLE OF INVENTION:  GASTROINTESTINAL TRACT DISORDERS
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Amgen Inc.
;      STREET:  1840 Dehavilland Drive
;      CITY:  Thousand Oaks
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  91320
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/622,720A
;      FILING DATE:
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Winter, Robert B.
;      REFERENCE/DOCKET NUMBER:  A-396
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-622-720A-23


  Query Match             29.4%;  Score 30;  DB 2;  Length 17;
  Best Local Similarity   53.8%;  Pred. No. 21;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVTV 19
           |||| |  :  ||
Db       1 VVGGQPGNSPWTV 13


RESULT  10
US-08-943-363-112
; Sequence 112, Application US/08943363
; Patent No. 5837478
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Van der Vieren, Monica
;    TITLE OF INVENTION:  No. 5837478el Human  2 Integrin Alpha Subunit
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive, 6300 Sear Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,363
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/173,497
;      FILING DATE:  23-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,889
;      FILING DATE:  5-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/362,652
;      FILING DATE:  21-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams Jr., Joseph A.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  27866/32684
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  112:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-943-363-112


  Query Match             28.4%;  Score 29;  DB 2;  Length 17;
  Best Local Similarity   75.0%;  Pred. No. 32;
  Matches    6;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 IFQEVVGG 10
             ||||  ||
Db      10 IFQEDAGG 17


RESULT  11
US-08-666-473-114
; Sequence 114, Application US/08666473
; Patent No. 5843713
;  GENERAL INFORMATION:
;    APPLICANT:  YOSHIDA, Aruto
;    APPLICANT:  TAKEUCHI, Makoto
;    TITLE OF INVENTION:  PEPTIDE SEQUENCE THAT FORMS MUCIN SUGAR
;    TITLE OF INVENTION:  CHAIN AND TECHNIQUE FOR MODIFYING PROTEIN TO BE LINKED
;    TITLE OF INVENTION:  WITH MUCIN SUGAR CHAIN
;    NUMBER OF SEQUENCES:  114
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/666,473
;      FILING DATE:  19-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/JP95/02238
;      FILING DATE:  01-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-22101
;      FILING DATE:  09-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 6-269111
;      FILING DATE:  01-NOV-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/837
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  114:
```

```
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/486,348A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C6
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-486-348A-41


  Query Match             30.4%;  Score 31;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   7
US-08-468-545B-41
; Sequence 41, Application US/08468545B
; Patent No. 5876712
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,545B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C5
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-468-545B-41


  Query Match             30.4%;  Score 31;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   8
US-08-466-680B-41
; Sequence 41, Application US/08466680B
; Patent No. 6075122
;  GENERAL INFORMATION:
;    APPLICANT:  Cheever, Martin A.
;    APPLICANT:  Disis, Mary L.
;    TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;    TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;    TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;    NUMBER OF SEQUENCES:  69
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,680B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sharkey, Richard G.
;      REGISTRATION NUMBER:  32,629
;      REFERENCE/DOCKET NUMBER:  920010.448C4
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-466-680B-41


  Query Match             30.4%;  Score 31;  DB 3;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 GGFPSQAQVTV 19
           |  |:|||: :
Db       1 GAMPNQAQMRI 11


RESULT   9
US-08-622-720A-23
; Sequence 23, Application US/08622720A
; Patent No. 5814308
;  GENERAL INFORMATION:
;    APPLICANT:  Zhang, Ke
```

```
Db      1 VVGGXPGNSPXTV 13


RESULT   3
5219991-8
;Patent No. 5219991
;     APPLICANT: LEONARD, EDWARD;SKEEL, ALISON H.;YOSHIMURA.
;TEIZO;APPELLA, ETTORE
;     TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN
;     NUMBER OF SEQUENCES: 12
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/586,085
;       FILING DATE: 21-SEP-1990
;SEQ ID NO:8:
;       LENGTH: 20
5219991-8


  Query Match             31.4%;  Score 32;  DB 5;  Length 20;
  Best Local Similarity   53.8%;  Pred. No. 11;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      7 VVGGFPSQAQVTV 19
          |||| |   :  ||
Db      1 VVGGHPGNSPWTV 13


RESULT   4
US-08-467-083-41
; Sequence 41, Application US/08467083
; Patent No. 5726023
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/NEU PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/NEU ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  68
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry
;       STREET:  6300 Columbia Center, 701 Fifth Avenue
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  US
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/467,083
;       FILING DATE:  06-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/414,417
;       FILING DATE:  06-JUN-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;       TELEX:  3723836 SEEDANBERRY
;  INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-467-083-41


  Query Match             30.4%;  Score 31;  DB 1;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      9 GGFPSQAQVTV 19
          |  |:|||: :
Db      1 GAMPNQAQMRI 11


RESULT   5
US-08-414-417B-41
; Sequence 41, Application US/08414417B
; Patent No. 5801005
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  69
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry LLP
;       STREET:  6300 Columbia Center, 701 Fifth Avenue
;       CITY:  Seattle
;       STATE:  Washington
;       COUNTRY:  US
;       ZIP:  98104-7092
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/414,417B
;       FILING DATE:  31-MAR-1995
;       CLASSIFICATION:  424
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sharkey, Richard G.
;       REGISTRATION NUMBER:  32,629
;       REFERENCE/DOCKET NUMBER:  920010.448C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  41:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-414-417B-41


  Query Match             30.4%;  Score 31;  DB 1;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      9 GGFPSQAQVTV 19
          |  |:|||: :
Db      1 GAMPNQAQMRI 11


RESULT   6
US-08-486-348A-41
; Sequence 41, Application US/08486348A
; Patent No. 5846538
;  GENERAL INFORMATION:
;     APPLICANT:  Cheever, Martin A.
;     APPLICANT:  Disis, Mary L.
;     TITLE OF INVENTION:  IMMUNE REACTIVITY TO HER-2/neu PROTEIN
;     TITLE OF INVENTION:  FOR DIAGNOSIS AND TREATMENT OF MALIGNANCIES IN WHICH THE
;     TITLE OF INVENTION:  HER-2/neu ONCOGENE IS ASSOCIATED
;     NUMBER OF SEQUENCES:  69
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Seed and Berry LLP
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:47 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-8
Perfect score: 102
Sequence: 1 QEIFQEVVGGFPSQAQVTVH 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	36	35.3	14	3	US-09-120-365-87	Sequence 87, Appl
2	34	33.3	19	5	5219991-12	Patent No. 5219991
3	32	31.4	20	5	5219991-8	Patent No. 5219991
4	31	30.4	15	1	US-08-467-083-41	Sequence 41, Appl
5	31	30.4	15	1	US-08-414-417B-41	Sequence 41, Appl
6	31	30.4	15	2	US-08-486-348A-41	Sequence 41, Appl
7	31	30.4	15	2	US-08-468-545B-41	Sequence 41, Appl
8	31	30.4	15	3	US-08-466-680B-41	Sequence 41, Appl
9	30	29.4	17	2	US-08-622-720A-23	Sequence 23, Appl
10	29	28.4	17	2	US-08-943-363-112	Sequence 112, App
11	28	27.5	12	2	US-08-666-473-114	Sequence 114, App
12	27	26.5	8	2	US-08-669-284B-26	Sequence 26, Appl
13	27	26.5	11	1	US-08-173-497-15	Sequence 15, Appl
14	27	26.5	11	1	US-08-286-889-15	Sequence 15, Appl
15	27	26.5	11	1	US-08-485-618-15	Sequence 15, Appl
16	27	26.5	11	1	US-08-362-652-15	Sequence 15, Appl
17	27	26.5	11	2	US-08-605-672-15	Sequence 15, Appl
18	27	26.5	11	2	US-08-482-293A-15	Sequence 15, Appl
19	27	26.5	11	2	US-08-943-363-15	Sequence 15, Appl
20	27	26.5	14	3	US-09-120-365-85	Sequence 85, Appl
21	27	26.5	16	2	US-08-474-696A-7	Sequence 7, Appli
22	27	26.5	17	3	US-08-256-747C-22	Sequence 22, Appl
23	27	26.5	19	2	US-08-474-696A-1	Sequence 1, Appli
24	27	26.5	19	2	US-08-474-696A-3	Sequence 3, Appli
25	27	26.5	20	1	US-08-221-816B-11	Sequence 11, Appl
26	26	25.5	6	1	US-07-890-422B-15	Sequence 15, Appl
27	26	25.5	10	1	US-07-890-422B-19	Sequence 19, Appl
28	26	25.5	11	1	US-08-081-539-106	Sequence 106, App
29	26	25.5	11	1	US-08-466-647-106	Sequence 106, App
30	26	25.5	13	3	US-08-834-314-5	Sequence 5, Appli
31	26	25.5	14	3	US-09-120-365-82	Sequence 82, Appl
32	26	25.5	14	3	US-09-120-365-83	Sequence 83, Appl
33	26	25.5	14	3	US-09-120-365-84	Sequence 84, Appl
34	26	25.5	16	2	US-08-480-190-175	Sequence 175, App
35	26	25.5	16	2	US-08-488-379-175	Sequence 175, App
36	26	25.5	16	4	PCT-US93-07545-175	Sequence 175, App
37	25.5	25.0	20	1	US-08-218-025A-63	Sequence 63, Appl
38	25	24.5	5	3	US-08-570-761-3	Sequence 3, Appli
39	25	24.5	6	1	US-07-890-422B-14	Sequence 14, Appl
40	25	24.5	6	3	US-08-570-761-2	Sequence 2, Appli
41	25	24.5	7	1	US-08-482-880-32	Sequence 32, Appl
42	25	24.5	7	2	US-08-273-274-32	Sequence 32, Appl
43	25	24.5	7	2	US-08-475-041-32	Sequence 32, Appl
44	25	24.5	7	2	US-08-484-773-32	Sequence 32, Appl
45	25	24.5	9	2	US-08-934-222-89	Sequence 89, Appl

ALIGNMENTS

```
RESULT 1
US-09-120-365-87
; Sequence 87, Application US/09120365
; Patent No. 6103514
; GENERAL INFORMATION:
;  APPLICANT: Natori, Shunji
;  TITLE OF INVENTION: NEW PROTEASE
;  FILE REFERENCE: 32290-144749
;  CURRENT APPLICATION NUMBER: US/09/120,365
;  CURRENT FILING DATE: 1998-07-22
;  EARLIER APPLICATION NUMBER: JP 9-333 474
;  EARLIER FILING DATE: 1997-11-18
;  NUMBER OF SEQ ID NOS: 101
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 87
;   LENGTH: 14
;   TYPE: PRT
;   ORGANISM: Nicotiana
US-09-120-365-87


  Query Match             35.3%;  Score 36;  DB 3;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 1.3;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 EIFQEVVGGFPS 13
           ||  ||| | ||
Db       3 EIEDEVVAGLPS 14


RESULT 2
5219991-12
;Patent No. 5219991
;    APPLICANT: LEONARD, EDWARD;SKEEL, ALISON H.;YOSHIMURA.
;TEIZO;APPELLA, ETTORE
;    TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN
;    NUMBER OF SEQUENCES: 12
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/586,085
;      FILING DATE: 21-SEP-1990
;SEQ ID NO:12:
;      LENGTH: 19
5219991-12


  Query Match             33.3%;  Score 34;  DB 5;  Length 19;
  Best Local Similarity   53.8%;  Pred. No. 4.4;
  Matches    7;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 VVGGFPSQAQVTV 19
           |||| |  :  ||
```

Search completed: December 21, 2000, 08:35:38
Job time: 601 sec

```
RT   "Proteus mirabilis flagella and MR/P fimbriae: isolation,
RT   purification, N-terminal analysis, and serum antibody response
RT   following experimental urinary tract infection.";
RL   Infect. Immun. 59:3574-3580(1991).
SQ   SEQUENCE   20 AA;  1923 MW;  6B3AD0D464F4CD11 CRC64;


  Query Match             22.5%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   33.3%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 SGERAVMLLGRV 14
          ||:  :   |:|
Db      2 SGDGTITFTGKV 13


RESULT  12
Q9TWL3
ID   Q9TWL3      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWL3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   INITIATORIN (FRAGMENT).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95218779.
RA   Aigaki T., Kasuga H., Nagaoka S., Osanai M.;
RT   "Purification and partial amino acid sequence of initiatorin, a
RT   prostatic endopeptidase of the silkworm, Bombyx mori.";
RL   Insect Biochem. Mol. Biol. 24:969-975(1994).
DR   INTERPRO; IPR001254; -.
DR   PFAM; PF00089; trypsin; 1.
SQ   SEQUENCE   20 AA;  2205 MW;  8BE047E96CFFF1BA CRC64;


  Query Match             22.5%;  Score 23;  DB 5;  Length 20;
  Best Local Similarity   71.4%;  Pred. No. 3.3e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 VSGERAV 8
          | | |||
Db      2 VGGRRAV 8


RESULT  13
Q28734
ID   Q28734      PRELIMINARY;      PRT;    20 AA.
AC   Q28734;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   TROPONIN T (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 83167564.
RA   Putney S.D., Herlihy W.C., Schimmel P.R.;
RT   "A new troponin T and cDNA clones for 13 different muscle proteins,
RT   found by shotgun sequencing.";
RL   Nature 302:718-721(1983).
DR   EMBL; V00901; CAA24266.1; -.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2495 MW;  83572E605F56A884 CRC64;


  Query Match             22.5%;  Score 23;  DB 6;  Length 20;
  Best Local Similarity   41.2%;  Pred. No. 3.3e+03;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    4;  Gaps    1;

Qy      4 GERAVMLLGRVKYGLHN 20
          ||:    | | || : |
Db      3 GEK----LKRQKYDITN 15


RESULT  14
Q05403
ID   Q05403      PRELIMINARY;      PRT;     8 AA.
AC   Q05403;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; SaccharomyceLales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FY1679;
RX   MEDLINE; 96021609.
RA   Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT   "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT   more than twice as many unknown as known open reading frames.";
RL   Yeast 11:975-986(1995).
DR   EMBL; X83121; CAA58183.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  879 MW;  7B5322D2C441E058 CRC64;


  Query Match             21.6%;  Score 22;  DB 3;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     17 GLHN 20
          |:||
Db      1 GIHN 4


RESULT  15
P70007
ID   P70007      PRELIMINARY;      PRT;    14 AA.
AC   P70007;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HISTONE H4-1 (FRAGMENT).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85160855.
RA   Gargiulo G., Razvi F., Ruberti I., Mohr I., Worcel A.;
RT   "Chromatin-specific hypersensitive sites are assembled on a Xenopus
RT   histone gene injected into Xenopus oocytes.";
RL   J. Mol. Biol. 181:333-349(1985).
DR   EMBL; M23777; AAA49737.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   14 AA;  1524 MW;  65A76B0A927B34B4 CRC64;


  Query Match             21.6%;  Score 22;  DB 13;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GRVKYG 17
          ||  ||
Db      6 GRTLYG 11
```

```
RESULT   7
Q9TWT8
ID   Q9TWT8      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWT8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GLUTATHIONE TRANSFERASE ISOZYME II (EC 2.5.1.18) (FRAGMENT).
OS   Penaeus japonicus (Kuruma prawn).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC   Penaeidae; Penaeus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93278989.
RA   Lin K.S., Chuang N.N.;
RT   "Anionic glutathione S-transferases in shrimp eyes.";
RL   Comp. Biochem. Physiol. B, Comp. Biochem. 105:151-156(1993).
SQ   SEQUENCE   20 AA;  2254 MW;  5D6A7ABCF09AD338 CRC64;


  Query Match             23.5%;  Score 24;  DB 5;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 2.3e+03;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 ERAVML 10
          ||:|||
Db     13 ERSVML 18


RESULT   8
Q25355
ID   Q25355      PRELIMINARY;      PRT;    10 AA.
AC   Q25355;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   VITELLOGENIN A (FRAGMENT).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Oedipodinae; Locusta.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 88003971.
RA   Locke J., White B.N., Wyatt G.R.;
RT   "Cloning and 5' end nucleotide sequences of two juvenile hormone-
RT   inducible vitellogenin genes of the African migratory locust.";
RL   DNA 6:331-342(1987).
DR   EMBL; M17333; AAA29284.1; -.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1128 MW;  D1B31177272042CD CRC64;


  Query Match             22.5%;  Score 23;  DB 5;  Length 10;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 AVMLLG 12
          ||:|||
Db      3 AVILLG 8


RESULT   9
O24445
ID   O24445      PRELIMINARY;      PRT;    17 AA.
AC   O24445;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   THIS ORF RESIDES IN THE 5'UTR OF CPPHY2.
OS   Ceratodon purpureus (Moss).
OC   Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida;
OC   Dicranidae; Dicranales; Ditrichaceae; Ceratodon.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WT3;
RA   Hughes J.E., Lamparter T., Mittmann F.;
RL   Plant Physiol. 112:446-446(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WT3;
RA   Pasentsis K., Paulo N., Dittrich P., Algarra P., Thuemmler F.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U72993; AAB19059.1; -.
SQ   SEQUENCE   17 AA;  1949 MW;  072DF72059DF1C7C CRC64;


  Query Match             22.5%;  Score 23;  DB 10;  Length 17;
  Best Local Similarity   40.0%;  Pred. No. 2.8e+03;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 SGERAVMLLG 12
            |  |::|::|
Db      6 STSRSLMIVG 15


RESULT  10
Q53545
ID   Q53545      PRELIMINARY;      PRT;    19 AA.
AC   Q53545;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   VIRF (FRAGMENT).
GN   VIRF.
OS   Shigella sonnei.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Shigella.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95394812.
RA   Nakayama S., Watanabe H.;
RT   "Involvement of cpxA, a sensor of a two-component regulatory system,
RT   in the pH-dependent regulation of expression of Shigella sonnei virF
RT   gene.";
RL   J. Bacteriol. 177:5062-5069(1995).
DR   EMBL; S79443; AAB35192.1; -.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2344 MW;  52A53B9DAAEEFCF4 CRC64;


  Query Match             22.5%;  Score 23;  DB 2;  Length 19;
  Best Local Similarity   57.1%;  Pred. No. 3.1e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     14 VKYGLHN 20
          :|  |||
Db     12 IKVRLHN 18


RESULT  11
Q9R5T4
ID   Q9R5T4      PRELIMINARY;      PRT;    20 AA.
AC   Q9R5T4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   MRPA=MANNOSE-RESISTANT/PROTEUSLIKE (MR/P) FIMBRIAL SUBUNIT.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Proteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 91372967.
RA   Bahrani F.K., Johnson D.E., Robbins D., Mobley H.L.;
```

```
RX   MEDLINE; 95232503.
RA   Smith A.B., Esko J.D., Hajduk S.L.;
RT   "Killing of trypanosomes by the human haptoglobin-related protein.";
RL   Science 268:284-286(1995).
SQ   SEQUENCE   19 AA;  2072 MW;  5BE2440A3DB3ABD6 CRC64;


  Query Match             24.5%;  Score 25;  DB 5;  Length 19;
  Best Local Similarity   26.3%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 DVSGERAVMLLGRVKYGLH 19
          |:    :::| ||  : |:
Db      1 DLGAVISLLLXGRQLFALY 19


RESULT   3
Q15892
ID   Q15892      PRELIMINARY;      PRT;     9 AA.
AC   Q15892;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   (CLONE XP3B4A) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=PLACENTA;
RA   Lee C.C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.Y.,
RA   Caskey C.T.H.;
RL   Hum. Mol. Genet. 0:0-0(0).
DR   EMBL; L32071; AAA73882.1; -.
FT   NON_TER       1       1
FT   NON_TER       9       9
SQ   SEQUENCE   9 AA;  971 MW;  49B22732CDC40B17 CRC64;


  Query Match             23.5%;  Score 24;  DB 4;  Length 9;
  Best Local Similarity   83.3%;  Pred. No. 3e+05;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 ERAVML 10
          ||||:|
Db      3 ERAVLL 8


RESULT   4
Q9TQZ5
ID   Q9TQZ5      PRELIMINARY;      PRT;     17 AA.
AC   Q9TQZ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96273610.
RA   Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA   Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT   "Studies on the mechanism of early onset macular degeneration in
RT   cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT   of two proteins in the retina.";
RL   Exp. Eye Res. 62:211-219(1996).
SQ   SEQUENCE   17 AA;  1671 MW;  52CCD0D1A98B3DAF CRC64;


  Query Match             23.5%;  Score 24;  DB 6;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     12 GRVKYGLH 19
          |:|| |::
Db      1 GKVKVGVN 8


RESULT   5
Q57012
ID   Q57012      PRELIMINARY;      PRT;     19 AA.
AC   Q57012;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   LEADER PEPTIDE.
GN   ERMA.
OS   Staphylococcus aureus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 86135972.
RA   Murphy E., Huwyler L., De Freire Bastos M.D.C.;
RT   "Transposon Tn554: complete nucleotide sequence and isolation of
RT   transposition-defective and antibiotic-sensitive mutants.";
RL   EMBO J. 4:3357-3365(1985).
DR   EMBL; X03216; CAA26965.1; -.
SQ   SEQUENCE   19 AA;  2257 MW;  19F81AD99E4F2F9B CRC64;


  Query Match             23.5%;  Score 24;  DB 2;  Length 19;
  Best Local Similarity   23.1%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 GERAVMLLGRVKY 16
          |  :: :: :|:|
Db      2 GTFSIFVINKVRY 14


RESULT   6
Q9S8G6
ID   Q9S8G6      PRELIMINARY;      PRT;     19 AA.
AC   Q9S8G6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   TYPE 2 LECTIN (FRAGMENT).
OS   Colocasia esculenta (Elephant's ear) (Taro).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Araceae; Colocasia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95288362.
RA   Van Damme E.J., Goossens K., Smeets K., van Leuven F., Verhaert P.,
RA   Peumans W.J.;
RT   "The major tuber storage protein of araceae species is a lectin.
RT   Characterization and molecular cloning of the lectin from Arum
RT   maculatum L.";
RL   Plant Physiol. 107:1147-1158(1995).
SQ   SEQUENCE   19 AA;  2066 MW;  F9C18865CA58608A CRC64;


  Query Match             23.5%;  Score 24;  DB 10;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      8 VMLLGRVKYG 17
          ::  |:| ||
Db      8 LLFSGQVXYG 17
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:37 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-7
Perfect score: 102
Sequence: 1 DVSGERAVMLLGRVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	24.5	19	2	O31045	O31045 streptomyce
2	25	24.5	19	5	Q9TWK7	Q9twk7 trypanosoma
3	24	23.5	9	4	Q15892	Q15892 homo sapien
4	24	23.5	17	6	Q9TQZ5	Q9tqz5 macaca fasc
5	24	23.5	19	2	Q57012	Q57012 staphylococ
6	24	23.5	19	10	Q9S8G6	Q9s8g6 colocasia e
7	24	23.5	20	5	Q9TWT8	Q9twt8 penaeus jap
8	23	22.5	10	5	Q25355	Q25355 locusta mig
9	23	22.5	17	10	O24445	O24445 ceratodon p
10	23	22.5	19	2	Q53545	Q53545 shigella so
11	23	22.5	20	2	Q9R5T4	Q9r5t4 proteus mir
12	23	22.5	20	5	Q9TWL3	Q9twl3 bombyx mori
13	23	22.5	20	6	Q28734	Q28734 oryctolagus
14	22	21.6	8	3	Q05403	Q05403 saccharomyc
15	22	21.6	14	13	P70007	P70007 xenopus lae
16	22	21.6	15	2	Q9R4P5	Q9r4p5 brevundimon
17	22	21.6	15	6	Q9TQQ5	Q9tqq5 bos taurus
18	22	21.6	19	2	Q9R4X3	Q9r4x3 bacillus ce
19	22	21.6	20	6	Q9TRQ4	Q9trq4 bos taurus
20	22	21.6	20	10	Q9S739	Q9s739 arabidopsis
21	21	20.6	8	2	Q9R4M3	Q9r4m3 enterococcu
22	21	20.6	14	2	Q60199	Q60199 escherichia
23	21	20.6	14	12	Q85718	Q85718 reovirus sp
24	21	20.6	15	3	Q9URC5	Q9urc5 saccharomyc
25	21	20.6	15	6	Q9TRR9	Q9trr9 oryctolagus
26	21	20.6	15	12	Q79359	Q79359 human immun
27	21	20.6	15	12	Q97098	Q97098 human immun
28	21	20.6	17	4	Q9UD18	Q9ud18 homo sapien
29	21	20.6	17	10	Q9S8U7	Q9s8u7 solanum tub
30	21	20.6	17	11	O88868	O88868 mus musculu
31	21	20.6	17	11	Q9QY07	Q9qy07 mus musculu
32	21	20.6	19	2	Q9R2F7	Q9r2f7 escherichia
33	21	20.6	19	4	Q9UC73	Q9uc73 homo sapien
34	21	20.6	19	8	Q36277	Q36277 zea mays (m
35	21	20.6	19	11	Q9QXP9	Q9qxp9 mus musculu
36	21	20.6	19	12	Q84863	Q84863 unidentifie
37	21	20.6	19	12	Q85671	Q85671 reovirus sp
38	21	20.6	20	2	O34197	O34197 rickettsia
39	21	20.6	20	2	Q9R9A5	Q9r9a5 nitrosospir
40	21	20.6	20	2	Q9R987	Q9r987 nitrosospir
41	21	20.6	20	12	Q85670	Q85670 reovirus sp
42	21	20.6	20	13	Q9PRP2	Q9prp2 gallus gall
43	20	19.6	13	2	O34770	O34770 borrelia af
44	20	19.6	14	2	Q45876	Q45876 clostridium
45	20	19.6	15	10	P93515	P93515 arabidopsis

ALIGNMENTS

RESULT 1
O31045
ID O31045 PRELIMINARY; PRT; 19 AA.
AC O31045;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE LEUA LEADER PEPTIDE.
GN LEUA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D132;
RA Potter C.A., Baumberg S.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF026444; AAB82585.1; -.
SQ SEQUENCE 19 AA; 2080 MW; 1A591DC2999760D4 CRC64;

Query Match 24.5%; Score 25; DB 2; Length 19;
Best Local Similarity 44.4%; Pred. No. 1.5e+03;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

```
Qy      10 LLGRVKYGL 18
           :|| :::||
Db       1 MLGIMRFGL 9
```

RESULT 2
Q9TWK7
ID Q9TWK7 PRELIMINARY; PRT; 19 AA.
AC Q9TWK7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HAPTOGLOBIN-RELATED PROTEIN ALPHA SUBUNIT (FRAGMENT).
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
RN [1]
RP SEQUENCE.

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4359;
RA   Wilson S.D., Wang M., Filpula D.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE
CC       AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC       ARGININE DEIMINASE PATHWAY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC       PERMEASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07185; AAA16963.1; -.
KW   Transport; Amino-acid transport; Transmembrane; Inner membrane.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1644 MW;  90B48A7C8FAA9705 CRC64;


  Query Match             19.6%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLHN 20
             |||::
Db       7 YGLYD 11


RESULT  14
PGK_VIBCH
ID   PGK_VIBCH      STANDARD;      PRT;    16 AA.
AC   P96154;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (FRAGMENT).
GN   PGK.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=OGAWA 395;
RX   MEDLINE; 97136637.
RA   Carroll P.A., Zhao G., Boyko S.A., Winkler M.E., Calderwood S.B.;
RT   "Identification, sequencing, and enzymatic activity of the
RT   erythrose-4-phosphate dehydrogenase gene of Vibrio cholerae.";
RL   J. Bacteriol. 179:293-296(1997).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +
CC       3-PHOSPHO-D-GLYCEROYL PHOSPHATE.
CC   -!- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U72152; AAC44768.1; -.
DR   INTERPRO; IPR001576; -.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; PARTIAL.
KW   Transferase; Kinase; Glycolysis.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1789 MW;  6BEA715980B8E227 CRC64;


  Query Match             19.6%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 DVSGER 6
             |::|:|
Db      10 DLAGKR 15


RESULT  15
COXQ_ONCMY
ID   COXQ_ONCMY     STANDARD;      PRT;    20 AA.
AC   P80335;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-1994 (Rel. 30, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIII-LIVER/HEART (EC 1.9.3.1)
DE   (FRAGMENT).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART, AND LIVER;
RX   MEDLINE; 94237150.
RA   Freund R., Kadenbach B.;
RT   "Identification of tissue-specific isoforms for subunits Vb and VIIa
RT   of cytochrome c oxidase isolated from rainbow trout.";
RL   Eur. J. Biochem. 221:1111-1116(1994).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
DR   PIR; S43635; S43635.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2048 MW;  4977B9DE80E62C58 CRC64;


  Query Match             19.6%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity   42.9%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 GERAVML 10
             ||:|: :
Db      13 GEQAIAM 19


Search completed: December 21, 2000, 08:32:57
Job time: 440 sec
```

```
CC   -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC       OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC   -!- TISSUE SPECIFICITY: MIDGUT.
CC   -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES      19      19       AMIDATION.
SQ   SEQUENCE   19 AA;  1930 MW;  99B5471A011625E5 CRC64;


  Query Match            20.6%;  Score 21;  DB 1;  Length 19;
  Best Local Similarity  80.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 SGERA 7
           :||||
Db       1 NGERA 5


RESULT  10
UP25_UPEIN
ID   UP25_UPEIN     STANDARD;      PRT;     19 AA.
AC   P82031;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.5.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST M.LUTEUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1940; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1941 MW;  5E94C6C757B463D9 CRC64;


  Query Match            20.6%;  Score 21;  DB 1;  Length 19;
  Best Local Similarity  50.0%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 LLGRVK 15
           :||::|
Db       9 VLGKIK 14


RESULT  11
COXQ_THUOB
ID   COXQ_THUOB     STANDARD;      PRT;     20 AA.
AC   P80983;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIII-1 (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART, AND LIVER;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20      20
SQ   SEQUENCE   20 AA;  1993 MW;  5BFEA5080A862C58 CRC64;


  Query Match            20.6%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity  37.5%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 SGERAVML 10
           :||:|: :
Db      12 AGEQAIAM 19


RESULT  12
AL16_CARMA
ID   AL16_CARMA     STANDARD;      PRT;      8 AA.
AC   P81819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 16.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES      8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  813 MW;  7C286B45AB476878 CRC64;


  Query Match            19.6%;  Score 20;  DB 1;  Length 8;
  Best Local Similarity  57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 GRVKYGL 18
           |   |||
Db       2 GPYSYGL 8


RESULT  13
ARCD_PSEPU
ID   ARCD_PSEPU     STANDARD;      PRT;     16 AA.
AC   P41147;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   ARGININE/ORNITHINE ANTIPORTER (FRAGMENT).
GN   ARCD.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
RN   [1]
```

```
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN F.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES      14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1441 MW;  4D1653612B9DECC3 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   44.4%;  Pred. No. 9e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGRVKYGL 18
           |:|:|  |:
Db       4 LIGKVLSGI 12


RESULT   7
DSIP_RABIT
ID   DSIP_RABIT     STANDARD;      PRT;     9 AA.
AC   P01158;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   DELTA SLEEP-INDUCING PEPTIDE (DSIP).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 77185324.
RA   Monnier M., Dudler L., Gachter R., Maier P.F., Tobler H.J.,
RA   Schoenenberger G.A.;
RT   "The delta sleep inducing peptide (DSIP). Comparative properties of
RT   the original and synthetic nonapeptide.";
RL   Experientia 33:548-552(1977).
RN   [2]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE; 79054421.
RA   Schoenenberger G.A., Maier P.F., Tobler H.J., Wilson K., Monnier M.;
RT   "The delta EEG (sleep)-inducing peptide (DSIP). XI. Amino-acid
RT   analysis, sequence, synthesis and activity of the nonapeptide.";
RL   Pflugers Arch. 376:119-129(1978).
CC   -!- FUNCTION: WHEN INFUSED INTO THE MESODIENCEPHALIC VENTRICLE OF
CC       RECIPIENT RABBITS INDUCES SPINDLE AND DELTA EEG ACTIVITY AND
CC       REDUCED MOTOR ACTIVITIES.
CC   -!- MISCELLANEOUS: THIS PEPTIDE WAS OBTAINED FROM DIALYSATES OF
CC       OCCIPITAL VENOUS SINUS BLOOD FROM RABBITS KEPT ASLEEP BY ELECTRIC
CC       STIMULATION OF THE THALAMUS.
DR   PIR; A01422; QDRB.
SQ   SEQUENCE   9 AA;  849 MW;  DDD365BDDAA8787D CRC64;


  Query Match             20.6%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 DVSGE 5
             | |||
Db       5 DASGE 9


RESULT   8
PHLC_STAIN
ID   PHLC_STAIN     STANDARD;      PRT;     19 AA.
AC   P80924;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   SPHINGOMYELINASE C (EC 3.1.4.12) (BETA-TOXIN) (BETA-HEMOLYSIN)
DE   (NEUTRAL SPHINGOMYELINASE) (FRAGMENT).
OS   Staphylococcus intermedius.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Staphylococcus.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=94-072594;
RX   MEDLINE; 97072006.
RA   Dziewanowska K., Edwards V.M., Deringer J.R., Bohach G.A.,
RA   Guerra D.J.;
RT   "Comparison of the beta-toxins from Staphylococcus aureus and
RT   Staphylococcus intermedius.";
RL   Arch. Biochem. Biophys. 335:102-108(1996).
CC   -!- FUNCTION: REQUIRED, WITH SPHINGOMYELINASE TO EFFECT TARGET CELL
CC       LYSIS (HEMOLYSIS). THE PH OPTIMUM IS 6.0-7.5. IT HAS A HIGH
CC       SPECIFICITY FOR SPHINGOMYELIN, HYDROLYZES LYSOPHOSPHATIDYLCHOLINE
CC       AT A MUCH LOWER RATE, BUT HAS NO ACTIVITY TOWARDS
CC       PHOSPHATIDYLCHOLINE, PHOSPHATIDYLETHANOLAMINE, OR
CC       PHOSPHATIDYLSERINE.
CC   -!- CATALYTIC ACTIVITY: SPHINGOMYELIN + H(2)O = N-ACYLSPHINGOSINE +
CC       CHOLINE PHOSPHATE.
CC   -!- COFACTOR: MAGNESIUM.
CC   -!- SUBUNIT: MONOMER.
KW   Hemolysis; Hydrolase; Toxin; Magnesium.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2089 MW;  A50753FCF500F80E CRC64;


  Query Match             20.6%;  Score 21;  DB 1;  Length 19;
  Best Local Similarity   38.9%;  Pred. No. 1.8e+03;
  Matches    7;  Conservative    0;  Mismatches    7;  Indels    4;  Gaps    1;

Qy       3 SGERAVMLLGRVKYGLHN 20
           |||    |    |   ||
Db       2 SGENQAEL----KLATHN 15


RESULT   9
TRP3_LEUMA
ID   TRP3_LEUMA     STANDARD;      PRT;     19 AA.
AC   P81735;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TACHYKININ-RELATED PEPTIDE 3 (LEMTRP 3).
OS   Leucophaea maderae (Madeira cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Leucophaea.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=MIDGUT;
RX   MEDLINE; 97053012.
RA   Muren J.E., Naessel D.R.;
RT   "Isolation of five tachykinin-related peptides from the midgut of
RT   the cockroach Leucophaea madera: existence of N-terminally extended
RT   isoforms.";
RL   Regul. Pept. 65:185-196(1996).
```

```
RT   "Two-dimensional gel electrophoresis separation and N-terminal
RT   sequence analysis of proteins from Clostridium pasteurianum W5.";
RL   Electrophoresis 19:802-806(1998).
CC   -!- CATALYTIC ACTIVITY: ATP + 3-PHOSPHO-D-GLYCERATE = ADP +
CC       3-PHOSPHO-D-GLYCEROYL PHOSPHATE.
CC   -!- PATHWAY: SECOND STEP IN THE SECOND PHASE OF GLYCOLYSIS.
CC   -!- SUBUNIT: MONOMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN IS:
CC       5.6, ITS MW IS: 56.2 KDA.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
DR   INTERPRO; IPR001576; -.
DR   PROSITE; PS00111; PGLYCERATE_KINASE; PARTIAL.
KW   Transferase; Kinase; Glycolysis.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2355 MW;  749E31E595C85529 CRC64;


  Query Match             23.5%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 6e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 DVSGERAVM 9
            || |:| ::
Db      11 DVKGKRVLV 19


RESULT   3
TAL3_PICJA
ID   TAL3_PICJA     STANDARD;      PRT;     9 AA.
AC   P17441;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).
OS   Pichia jadinii (Yeast) (Candida utilis).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Pichia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 75145197.
RA   Tsolas O., Sun S.C.;
RT   "Isolation of a peptide containing a histidinyl-cysteinyl sequence
RT   from the active center of transaldolase.";
RL   Arch. Biochem. Biophys. 167:525-533(1975).
CC   -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
CC       METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC       3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
CC   -!- PATHWAY: NONOXYDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
DR   PIR; A11497; A11497.
DR   INTERPRO; IPR001585; -.
DR   PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR   PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW   Transferase; Pentose shunt.
FT   NON_TER      1      1
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  1033 MW;  325A31A44EB1E058 CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLH 19
           ||:|
Db       1 YGIH 4


RESULT   4
CRBL_VESOR
ID   CRBL_VESOR     STANDARD;      PRT;     14 AA.
```

```
AC   P17236;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1990 (Rel. 15, Last annotation update)
DE   HISTAMINE RELEASING PEPTIDE II (HR-II).
OS   Vespa orientalis (Oriental hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Miroshnikov A.I., Snezhkova L.G., Nazimov I.V., Reshetova O.I.,
RA   Rozynov B.V., Gushchin I.S.;
RT   "Structure and properties of histamine releasing peptides from the
RT   venom of Vespa orientalis hornet.";
RL   Bioorg. Khim. 7:1467-1477(1981).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
DR   PIR; JN0390; JN0390.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES      14     14       AMIDATION.
SQ   SEQUENCE   14 AA;  1524 MW;  22015B4A6CEDFD38 CRC64;


  Query Match             22.5%;  Score 23;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 6.1e+02;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 MLLGRVKYGL 18
            ::||::  ||
Db       4 LILGKLVKGL 13


RESULT   5
CRBL_VESTR
ID   CRBL_VESTR     STANDARD;      PRT;     13 AA.
AC   P17231;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VESPID CHEMOTACTIC PEPTIDE T (VESCP-T).
OS   Vespa tropica (Hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Yasuhara T., Nakajima T., Erspaer V.;
RL   (In) Sakakibara S. (eds.);
RL   Peptide chemistry 1982, pp.213-218, Protein Research Foundation,
RL   Osaka (1983).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES      13     13       AMIDATION.
SQ   SEQUENCE   13 AA;  1354 MW;  220140365DFE5338 CRC64;


  Query Match             21.6%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 8.4e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGRVKYGL 18
           :||::  ||
Db       4 ILGKILGGL 12


RESULT   6
TEMF_RANTE
ID   TEMF_RANTE     STANDARD;      PRT;     14 AA.
AC   P56921;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:56 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-7
Perfect score: 102
Sequence: 1 DVSGERAVMLLGRVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	24.5	14	1	TEMA_RANTE	P56917 rana tempor
2	24	23.5	20	1	PGK_CLOPA	P81346 clostridium
3	23	22.5	9	1	TAL3_PICJA	P17441 pichia jadi
4	23	22.5	14	1	CRBL_VESOR	P17236 vespa orien
5	22	21.6	13	1	CRBL_VESTR	P17231 vespa tropi
6	22	21.6	14	1	TEMF_RANTE	P56921 rana tempor
7	21	20.6	9	1	DSIP_RABIT	P01158 oryctolagus
8	21	20.6	19	1	PHLC_STAIN	P80924 staphylococ
9	21	20.6	19	1	TRP3_LEUMA	P81735 leucophaea
10	21	20.6	19	1	UP25_UPEIN	P82031 uperoleia i
11	21	20.6	20	1	COXQ_THUOB	P80983 thunnus obe
12	20	19.6	8	1	AL16_CARMA	P81819 carcinus ma
13	20	19.6	16	1	ARCD_PSEPU	P41147 pseudomonas
14	20	19.6	16	1	PGK_VIBCH	P96154 vibrio chol
15	20	19.6	20	1	COXQ_ONCMY	P80335 oncorhynchu
16	20	19.6	20	1	FIBR_PACLE	P81070 pacifastacu
17	19.5	19.1	20	1	TRY8_ECOLI	P33789 escherichia
18	19	18.6	8	1	AL17_CARMA	P81820 carcinus ma
19	19	18.6	13	1	CRBL_VESMA	P17232 vespa manda
20	19	18.6	13	1	IDHC_PIG	P20304 sus scrofa
21	19	18.6	14	1	DCMM_PSECF	P19914 pseudomonas
22	19	18.6	14	1	TEMC_RANTE	P56918 rana tempor
23	19	18.6	14	1	UC18_MAIZE	P80624 zea mays (m
24	19	18.6	15	1	LPL_THETH	P21234 thermus aqu
25	19	18.6	15	1	URE1_MORMO	P17337 morganella
26	19	18.6	17	1	BOL4_MEGPE	P07495 megabombus
27	19	18.6	18	1	LUXB_KRYAS	P18300 kryptophana
28	19	18.6	19	1	FIBA_BUBAR	P14443 bubalus arn
29	19	18.6	20	1	KORC_METTM	P80906 methanobact
30	19	18.6	20	1	LEC1_ARTIN	P18671 artocarpus
31	19	18.6	20	1	MDH_KIBAR	P19978 kibdelospor
32	19	18.6	20	1	MDH_MICGL	P19979 microtetras
33	18	17.6	9	1	AL11_CARMA	P81814 carcinus ma
34	18	17.6	9	1	TAL1_PICJA	P17440 pichia jadi
35	18	17.6	15	1	FIBA_SYNCA	P14463 syncerus ca
36	18	17.6	16	1	CFAB_BOVIN	P81187 bos taurus
37	18	17.6	17	1	NU4M_TRIRU	Q36834 trichophyto
38	18	17.6	19	1	RS19_SPICI	O31159 spiroplasma
39	18	17.6	20	1	CUDP_VERCH	P80406 verticilliu
40	18	17.6	20	1	SAMP_PLEPL	P20677 pleuronecte
41	17	16.7	10	1	GON1_ALLMI	P37041 alligator m
42	17	16.7	10	1	RL16_ACHLA	P29221 acholeplasm
43	17	16.7	10	1	TKN1_SCYCA	P08608 scyliorhinu
44	17	16.7	10	1	TKNB_RANCA	P22689 rana catesb
45	17	16.7	10	1	TKS1_AEDAE	P42634 aedes aegyp

ALIGNMENTS

RESULT 1
TEMA_RANTE
ID TEMA_RANTE STANDARD; PRT; 14 AA.
AC P56917;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TEMPORIN A.
OS Rana temporaria (European common frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN [1]
RP SEQUENCE, AND SYNTHESIS.
RC TISSUE=SKIN;
RX MEDLINE; 97175050.
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA Barra D.;
RT "Temporins, antimicrobial peptides from the European red frog Rana
RT temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC BACTERIA.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC FAMILY.
KW Amphibian skin; Antibiotic; Amidation; Multigene family.
FT MOD_RES 14 14 AMIDATION.
SQ SEQUENCE 14 AA; 1469 MW; 601653612B9DECD4 CRC64;

Query Match 24.5%; Score 25; DB 1; Length 14;
Best Local Similarity 55.6%; Pred. No. 2.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy      10 LLGRVKYGL 18
           |:|||  |:
Db       4 LIGRVLSGI 12
```

RESULT 2
PGK_CLOPA
ID PGK_CLOPA STANDARD; PRT; 20 AA.
AC P81346;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PUTATIVE PHOSPHOGLYCERATE KINASE (EC 2.7.2.3) (CP 11) (FRAGMENT).
GN PGK.
OS Clostridium pasteurianum.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE.
RC STRAIN=W5;
RX MEDLINE; 98291870.
RA Flengsrud R., Skjeldal L.;

Query Match 22.5%; Score 23; DB 2; Length 17;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      11 LGRVKYG 17
           |||: :|
Db      11 LGRLYFG 17
```

RESULT 14
S58129
hypothetical protein - moss (Ceratodon purpureus)
C;Species: Ceratodon purpureus
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 31-Oct-1997
C;Accession: S58129
R;Hughes, J.; Mittmann, F.
submitted to the EMBL Data Library, July 1995
A;Description: The moss Ceratodon purpureus contains and expresses a second, conventiona
A;Reference number: S58129
A;Accession: S58129
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-17 <HUG>
A;Cross-references: GB:U56698; EMBL:X89725; NID:g1314836; PID:g1322246

Query Match 22.5%; Score 23; DB 2; Length 17;
Best Local Similarity 40.0%; Pred. No. 1.7e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 SGERAVMLLG 12
           |  |::|::|
Db       6 STSRSLMIVG 15
```

RESULT 15
B43594
fimbrial protein mrpA - Proteus mirabilis (fragment)
C;Species: Proteus mirabilis
C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C;Accession: B43594
R;Bahrani, F.K.; Johnson, D.E.; Robbins, D.; Mobley, H.L.T.
Infect. Immun. 59, 3574-3580, 1991
A;Title: Proteus mirabilis flagella and MR/P fimbriae: isolation, purification, N-termin
A;Reference number: A43594; MUID:91372967
A;Accession: B43594
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <BAH>

Query Match 22.5%; Score 23; DB 2; Length 20;
Best Local Similarity 33.3%; Pred. No. 2e+03;
Matches 4; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy       3 SGERAVMLLGRV 14
           ||:  :   |:|
Db       2 SGDGTITFTGKV 13
```

Search completed: December 21, 2000, 08:30:06
Job time: 270 sec

C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; protein digestion; serine proteinase

Query Match 23.5%; Score 24; DB 2; Length 20;
Best Local Similarity 71.4%; Pred. No. 1.3e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 VSGERAV 8
|:|| ||
Db 2 VNGENAV 8

RESULT 9
S65399
immunodeficiency virus type 1, HIV-1 gp120 - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S65399
R;Niwa, Y.; Yano, M.; Futaki, S.; Okumura, Y.; Kido, H.
Eur. J. Biochem. 237, 64-70, 1996
A;Title: T-cell membrane-associated serine protease, tryptase TL(2), binds human immunod
man immunodeficiency virus type 1 inhibit cleavage of gp120.
A;Reference number: S65399; MUID:96203909
A;Accession: S65399
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10;11-20 <NIW>
C;Superfamily: type E retrovirus env polyprotein

Query Match 23.5%; Score 24; DB 2; Length 20;
Best Local Similarity 44.4%; Pred. No. 1.3e+03;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 6 RAVMLLGRV 14
| || : :||:
Db 5 RAFVTIGRI 13

RESULT 10
A11497
transaldolase (EC 2.2.1.2) III - yeast (Pichia jadinii) (fragment)
C;Species: Pichia jadinii, Candida utilis
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
C;Accession: A11497
R;Tsolas, O.; Sun, S.C.
Arch. Biochem. Biophys. 167, 525-533, 1975
A;Title: Isolation of a peptide containing a histidinyl-cysteinyl sequence from the acti
A;Reference number: A11497; MUID:75145197
A;Accession: A11497
A;Molecule type: protein
A;Residues: 1-9 <TSO>
C;Keywords: transferase

Query Match 22.5%; Score 23; DB 2; Length 9;
Best Local Similarity 75.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 16 YGLH 19
||:|
Db 1 YGIH 4

RESULT 11
S26557
T-cell receptor beta chain (clone Cw3/HLA2A3) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 13-Jan-1995 #sequence_revision 17-Apr-1998 #text_change 17-Mar-1999
C;Accession: S26557
R;Casanova, J.L.; Cerottini, J.C.; Matthes, M.; Necker, A.; Gournier, H.; Barra, C.; Wid
J. Exp. Med. 176, 439-447, 1992
A;Title: H-2-restricted cytolytic T lymphocytes specific for HLA display T cell recep
A;Reference number: S26512; MUID:92364546
A;Accession: S26557
A;Molecule type: mRNA
A;Residues: 1-12 <CAS>
A;Cross-references: EMBL:X68007
A;Experimental source: cytolytic T-lymphocyte, clone Cw3/HLA2A3
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 22.5%; Score 23; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 1.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 12 GRVKY 16
|||:|
Db 6 GRVEY 10

RESULT 12
JN0390
histamine-releasing peptide II - oriental hornet
N;Alternate names: venom protein HR-2
C;Species: Vespa orientalis (oriental hornet)
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 23-Aug-1997
C;Accession: JN0390; S10919
R;Miroshnikov, A.I.; Snezhkova, L.G.; Nazimov, I.V.; Reshetova, O.I.; Rozinov, B.V.;
Bioorg. Khim. 7, 1467-1477, 1981
A;Title: Structure and properties of histamine releasing peptides from the venom of V
A;Reference number: JN0389
A;Accession: JN0390
A;Molecule type: protein
A;Residues: 1-14 <MIR>
R;Tuichibaev, M.U.; Akhmedova, N.U.; Kazakov, I.; Korneev, A.S.; Gagel'gans, A.I.
Biochemistry (N.Y.) 53, 183-190, 1988
A;Title: Low-molecular-weight peptides of venom of the giant hornet Vespa orientalis.
A;Reference number: S06445
A;Accession: S10919
A;Molecule type: protein
A;Residues: 1-14 <TUI>
C;Superfamily: crabrolin
C;Keywords: amidated carboxyl end; venom
F;14/Modified site: amidated carboxyl end (Leu) #status experimental

Query Match 22.5%; Score 23; DB 2; Length 14;
Best Local Similarity 40.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 9 MLLGRVKYGL 18
::||:: ||
Db 4 LILGKLVKGL 13

RESULT 13
PH1754
T cell receptor alpha chain V region (clone 1V alpha 24-1) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 17-Mar-1999
C;Accession: PH1754
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1754
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-17 <POR>
C;Superfamily: immunoglobulin V region; immunoglobulin homology

```
RESULT   3
B56899
serum heterodimer, 24K chain - sandbar shark (fragment)
C;Species: Carcharhinus plumbeus (sandbar shark)
C;Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 20-Jun-2000
C;Accession: B56899
R;Vazquez-Moreno, L.; Porath, J.; Schluter, S.F.; Marchalonis, J.J.
Comp. Biochem. Physiol. B 103, 563-568, 1992
A;Title: Purification of a novel heterodimer from shark (Carcharhinus plumbeus) serum by
A;Reference number: A56899; MUID:93092592
A;Accession: B56899
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <VAZ>
C;Keywords: glycoprotein; plasma


  Query Match             23.5%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 6.6e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 VSGERAV 8
           | ||| |
Db       2 VGGERVV 8


RESULT   4
S74147
glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S74147
R;Fukuda, A.; Osawa, T.; Hitomi, K.; Uchida, K.
Arch. Biochem. Biophys. 333, 419-426, 1996
A;Title: 4-Hydroxy-2-nonenal cytotoxicity in renal proximal tubular cells: protein modif
A;Reference number: S74147; MUID:96404942
A;Accession: S74147
A;Molecule type: protein
A;Residues: 1-10 <FUK>
A;Experimental source: LLC-PK1 cells (renal tubular epithelial cells)
C;Keywords: NAD; oxidoreductase


  Query Match             23.5%;  Score 24;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           |:|| |::
Db       1 GKVKVGVN 8


RESULT   5
PU0025
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) - rice (strain Nihonbare) (fragment)
C;Species: Oryza sativa (rice)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 23-Mar-1995
C;Accession: PU0025
R;Tsugita, A.
submitted to JIPID, April 1993
A;Reference number: PS0206
A;Accession: PU0025
A;Molecule type: protein
A;Residues: 1-15 <TSU>
A;Experimental source: root
C;Keywords: oxidoreductase


  Query Match             23.5%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 1e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 VSGERAV 8
           :| ||||
Db       5 ISDERAV 11


RESULT   6
PH0779
T-cell receptor alpha chain (B83) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0779
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility comple
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0779
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60877
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor


  Query Match             23.5%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 LGRVKYG 17
           |||: :|
Db       9 LGRLHFG 15


RESULT   7
PH1455
T-cell receptor alpha chain (clone A24/PEF4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1455
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.;
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompa
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1455
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             23.5%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 LGRVKYG 17
           |||: :|
Db       9 LGRLHFG 15


RESULT   8
B61333
chymotrypsin (EC 3.4.21.1) - edible frog (fragment)
C;Species: Rana esculenta (edible frog)
C;Date: 17-Jul-1994 #sequence_revision 17-Jul-1994 #text_change 07-May-1999
C;Accession: B61333
R;Pies, W.; Zwilling, R.; Woodbury, R.G.; Neurath, H.
FEBS Lett. 109, 45-49, 1980
A;Title: Amino-terminal amino acid sequences and the evolution of frog (Rana esculent
A;Reference number: A61333; MUID:80113255
A;Accession: B61333
A;Molecule type: protein
A;Residues: 1-20 <PIE>
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:05 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-7
Perfect score: 102
Sequence: 1 DVSGERAVMLLGRVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	24.5	20	2	PH1338	Ig heavy chain DJ
2	24.5	24.0	15	2	PH0775	T-cell receptor al
3	24	23.5	10	2	B56899	serum heterodimer,
4	24	23.5	10	2	S74147	glyceraldehyde-3-p
5	24	23.5	15	2	PU0025	ubiquinol--cytochr
6	24	23.5	15	2	PH0779	T-cell receptor al
7	24	23.5	15	2	PH1455	T-cell receptor al
8	24	23.5	20	2	B61333	chymotrypsin (EC 3
9	24	23.5	20	2	S65399	immunodeficiency v
10	23	22.5	9	2	A11497	transaldolase (EC
11	23	22.5	12	2	S26557	T-cell receptor be
12	23	22.5	14	2	JN0390	histamine-releasin
13	23	22.5	17	2	PH1754	T cell receptor al
14	23	22.5	17	2	S58129	hypothetical prote
15	23	22.5	20	2	B43594	fimbrial protein m
16	22.5	22.1	15	2	PA0106	protein QF200076 -
17	22	21.6	14	2	I51432	histone H4-1 precu
18	22	21.6	15	2	A27504	histone H2A - mous
19	22	21.6	18	2	I52623	hypothetical prote
20	22	21.6	20	2	S32502	calpain (EC 3.4.22
21	22	21.6	20	2	S10876	hypothetical prote
22	21	20.6	9	2	QDRB	delta sleep-induci
23	21	20.6	12	2	S65629	protoporphyrinogen
24	21	20.6	15	2	A49887	transcription fact
25	21	20.6	16	2	S02455	DNA-invertase - ph
26	21	20.6	16	2	S65709	major allergen Myr
27	21	20.6	16	2	PH1475	T-cell receptor be
28	21	20.6	17	2	PH1802	T cell receptor al
29	21	20.6	17	2	S78421	ribosomal protein
30	21	20.6	18	2	A39997	group III allergen
31	21	20.6	19	2	E49048	T-cell receptor be
32	21	20.6	20	2	S77991	cytochrome-c oxida
33	21	20.6	20	2	PH1358	Ig heavy chain DJ
34	21	20.6	20	2	S78419	ribosomal protein
35	20	19.6	12	2	A61332	Na+/K+-exchanging
36	20	19.6	12	2	PH1635	Ig H chain V-D-J r
37	20	19.6	13	2	PS0277	glyceraldehyde-3-p
38	20	19.6	13	2	PS0325	tetrahydroberberin
39	20	19.6	14	2	S58866	botulinum neurotox
40	20	19.6	15	2	A30330	neuropeptide pep -
41	20	19.6	16	2	S36876	aquacobalamin redu
42	20	19.6	18	2	S54272	CTC 75 protein - h
43	20	19.6	18	2	G42753	interferon alpha (
44	20	19.6	19	2	A39504	octamer-binding pr
45	20	19.6	20	2	F42762	multicatalytic end

ALIGNMENTS

RESULT 1
PH1338
Ig heavy chain DJ region (clone C372-115) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1338
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor ly
A;Reference number: PH1302; MUID:93094761
A;Accession: PH1338
A;Molecule type: DNA
A;Residues: 1-20 <WAS>
C;Keywords: heterotetramer; immunoglobulin

Query Match 24.5%; Score 25; DB 2; Length 20;
Best Local Similarity 30.0%; Pred. No. 9e+02;
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy       8 VMLLGRVKYG 17
           :: :| ::||
Db       4 ILRIGLLRYG 13
```

RESULT 2
PH0775
T-cell receptor alpha chain (B28) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0775
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility comple
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0775
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60871
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 24.0%; Score 24.5; DB 2; Length 15;
Best Local Similarity 43.8%; Pred. No. 8.2e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 3; Gaps 1;

```
Qy       2 VSGERAVMLLGRVKYG 17
           :||  |   |||: :|
Db       3 LSGGSA---LGRLHFG 15
```

```
  Query Match             28.4%;  Score 29;  DB 2;  Length 16;
  Best Local Similarity   62.5%;  Pred. No. 53;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           ||:|| |:
Db       8 GRIKYTLN 15


RESULT  15
US-08-488-379-76
; Sequence 76, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,379
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  76:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-488-379-76


  Query Match             28.4%;  Score 29;  DB 2;  Length 16;
  Best Local Similarity   62.5%;  Pred. No. 53;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           ||:|| |:
Db       8 GRIKYTLN 15


Search completed: December 21, 2000, 08:31:47
Job time: 370 sec
```

```
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,379
;       FILING DATE:
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/077,255
;       FILING DATE:  June 15, 1993
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  77:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
US-08-488-379-77


  Query Match             28.4%;  Score 29;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 49;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           ||:|| |:
Db       8 GRIKYTLN 15


RESULT  13
PCT-US93-07545-77
; Sequence 77, Application PC/TUS9307545
;  GENERAL INFORMATION:
;     APPLICANT:  Robert G. Urban
;     APPLICANT:  Roman M. Chicz
;     APPLICANT:  Dario A. A. Vignali
;     APPLICANT:  Mary L. Hedley
;     APPLICANT:  Lawrence J. Stern
;     APPLICANT:  Jack L. Strominger
;     TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;     NUMBER OF SEQUENCES:  273
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/07545
;       FILING DATE:  19930811
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  77:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US93-07545-77


  Query Match             28.4%;  Score 29;  DB 4;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 49;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           ||:|| |:
Db       8 GRIKYTLN 15


RESULT  14
US-08-480-190-76
; Sequence 76, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;     APPLICANT:  Robert G. Urban
;     APPLICANT:  Roman M. Chicz
;     APPLICANT:  Dario A. A. Vignali
;     APPLICANT:  Mary L. Hedley
;     APPLICANT:  Lawrence J. Stern
;     APPLICANT:  Jack L. Strominger
;     TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;     NUMBER OF SEQUENCES:  274
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/480,190
;       FILING DATE:
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/077,255
;       FILING DATE:  June 15, 1993
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  76:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  16
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
US-08-480-190-76
```

```
;    MOLECULE TYPE:  peptide
US-08-164-618-20


  Query Match             28.4%;  Score 29;  DB 1;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 38;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 VKYGLHN 20
           ||:| ||
Db       3 VKHGSHN 9


RESULT  10
US-08-159-340A-17
; Sequence 17, Application US/08159340A
; Patent No. 5565352
;  GENERAL INFORMATION:
;    APPLICANT:  Hochstrasser, Mark
;    APPLICANT:  Papa, Feroz
;    TITLE OF INVENTION:  DEUBIQUITINATING ENZYME:  COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,340A
;      FILING DATE:  24-NOV-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  ARCD:112/HYL
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-340A-17


  Query Match             28.4%;  Score 29;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 45;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 VKYGLHN 20
           | || ||
Db       8 VHYGTHN 14


RESULT  11
US-08-480-190-77
; Sequence 77, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  77:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-77


  Query Match             28.4%;  Score 29;  DB 2;  Length 15;
  Best Local Similarity   62.5%;  Pred. No. 49;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GRVKYGLH 19
           ||:|| |:
Db       8 GRIKYTLN 15


RESULT  12
US-08-488-379-77
; Sequence 77, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
```

```
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/488,379
;       FILING DATE:
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/077,255
;       FILING DATE:  June 15, 1993
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  53:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
US-08-488-379-53


  Query Match             29.4%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   61.5%;  Pred. No. 45;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 DVSGERAVMLLGR 13
           ||   |||  |||
Db       1 DVGEYRAVTELGR 13


RESULT   8
PCT-US93-07545-53
; Sequence 53, Application PC/TUS9307545
;  GENERAL INFORMATION:
;     APPLICANT:  Robert G. Urban
;     APPLICANT:  Roman M. Chicz
;     APPLICANT:  Dario A. A. Vignali
;     APPLICANT:  Mary L. Hedley
;     APPLICANT:  Lawrence J. Stern
;     APPLICANT:  Jack L. Strominger
;     TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;     NUMBER OF SEQUENCES:  273
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  MS-DOS (Version 5.0)
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/07545
;       FILING DATE:  19930811
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  53:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US93-07545-53


  Query Match             29.4%;  Score 30;  DB 4;  Length 20;
  Best Local Similarity   61.5%;  Pred. No. 45;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 DVSGERAVMLLGR 13
           ||   |||  |||
Db       1 DVGEYRAVTELGR 13


RESULT   9
US-08-164-618-20
; Sequence 20, Application US/08164618
; Patent No. 5408036
;  GENERAL INFORMATION:
;     APPLICANT:  Ghadiri, M. Reza
;     TITLE OF INVENTION:  Isolated Metallopolypeptides:
;     TITLE OF INVENTION:  Compositions and Synthetic Methods
;     NUMBER OF SEQUENCES:  20
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Dressler, Goldsmith, Shore, Sutker &
;       ADDRESSEE:  Milnamow, Ltd.
;       STREET:  180 No. 5408036th Stetson, Suite 4700
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60601
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/164,618
;       FILING DATE:
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/769,621
;       FILING DATE:
;       APPLICATION NUMBER:  07/591,988
;       FILING DATE:  October 2, 1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Gamson, Edward G.
;       REGISTRATION NUMBER:  29,381
;       REFERENCE/DOCKET NUMBER:  SCRF 231.0
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 616-5400
;       TELEFAX:  (312) 616-5460
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  12 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
```

```
;       SOFTWARE:  WordPerfect (Version 5.1)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US93/07545
;       FILING DATE:  19930811
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/925,460
;       FILING DATE:  August 11, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00246/168001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  19
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
PCT-US93-07545-21


  Query Match             29.4%;  Score 30;  DB 4;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 42;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 DVSGERAVMLLGR 13
           ||   |||  |||
Db       1 DVGEYRAVTELGR 13


RESULT   5
PCT-US96-09303-21
; Sequence 21, Application PC/TUS9609303
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  IMMUNOGENIC PEPTIDES OF PROSTATE
;    TITLE OF INVENTION:  SPECIFIC ANTIGEN
;    NUMBER OF SEQUENCES:  22
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/09303
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/472,228
;      FILING DATE:  07-JUN-1995
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
PCT-US96-09303-21


  Query Match             29.4%;  Score 30;  DB 4;  Length 19;
  Best Local Similarity   75.0%;  Pred. No. 42;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 RAVMLLGR 13
           |:|:||||
Db      10 RSVILLGR 17


RESULT   6
US-08-480-190-53
; Sequence 53, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  53:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-53


  Query Match             29.4%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   61.5%;  Pred. No. 45;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 DVSGERAVMLLGR 13
           ||   |||  |||
Db       1 DVGEYRAVTELGR 13


RESULT   7
US-08-488-379-53
; Sequence 53, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
```

```
Db      10 RSVILLGR 17


RESULT   2
US-08-480-190-21
; Sequence 21, Application US/08480190
; Patent No. 5827516
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,190
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-480-190-21


  Query Match             29.4%;  Score 30;  DB 2;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 42;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 DVSGERAVMLLGR 13
          ||   |||  |||
Db      1 DVGEYRAVTELGR 13


RESULT   3
US-08-488-379-21
; Sequence 21, Application US/08488379
; Patent No. 5880103
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  274
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,379
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/077,255
;      FILING DATE:  June 15, 1993
;      APPLICATION NUMBER:  07/925,460
;      FILING DATE:  August 11, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00246/168001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-488-379-21


  Query Match             29.4%;  Score 30;  DB 2;  Length 19;
  Best Local Similarity   61.5%;  Pred. No. 42;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 DVSGERAVMLLGR 13
          ||   |||  |||
Db      1 DVGEYRAVTELGR 13


RESULT   4
PCT-US93-07545-21
; Sequence 21, Application PC/TUS9307545
;  GENERAL INFORMATION:
;    APPLICANT:  Robert G. Urban
;    APPLICANT:  Roman M. Chicz
;    APPLICANT:  Dario A. A. Vignali
;    APPLICANT:  Mary L. Hedley
;    APPLICANT:  Lawrence J. Stern
;    APPLICANT:  Jack L. Strominger
;    TITLE OF INVENTION:  IMMUNOMODULATORY PEPTIDES
;    NUMBER OF SEQUENCES:  273
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:46 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-7
Perfect score: 102
Sequence: 1 DVSGERAVMLLGRVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	29.4	19	1	US-08-472-228A-21	Sequence 21, Appl
2	30	29.4	19	2	US-08-480-190-21	Sequence 21, Appl
3	30	29.4	19	2	US-08-488-379-21	Sequence 21, Appl
4	30	29.4	19	4	PCT-US93-07545-21	Sequence 21, Appl
5	30	29.4	19	4	PCT-US96-09303-21	Sequence 21, Appl
6	30	29.4	20	2	US-08-480-190-53	Sequence 53, Appl
7	30	29.4	20	2	US-08-488-379-53	Sequence 53, Appl
8	30	29.4	20	4	PCT-US93-07545-53	Sequence 53, Appl
9	29	28.4	12	1	US-08-164-618-20	Sequence 20, Appl
10	29	28.4	14	1	US-08-159-340A-17	Sequence 17, Appl
11	29	28.4	15	2	US-08-480-190-77	Sequence 77, Appl
12	29	28.4	15	2	US-08-488-379-77	Sequence 77, Appl
13	29	28.4	15	4	PCT-US93-07545-77	Sequence 77, Appl
14	29	28.4	16	2	US-08-480-190-76	Sequence 76, Appl
15	29	28.4	16	2	US-08-488-379-76	Sequence 76, Appl
16	29	28.4	16	4	PCT-US93-07545-76	Sequence 76, Appl
17	29	28.4	18	2	US-08-480-190-192	Sequence 192, App
18	29	28.4	18	2	US-08-488-379-192	Sequence 192, App
19	29	28.4	18	4	PCT-US93-07545-192	Sequence 192, App
20	29	28.4	19	2	US-08-480-190-74	Sequence 74, Appl
21	29	28.4	19	2	US-08-480-190-75	Sequence 75, Appl
22	29	28.4	19	2	US-08-488-379-74	Sequence 74, Appl
23	29	28.4	19	2	US-08-488-379-75	Sequence 75, Appl
24	29	28.4	19	4	PCT-US93-07545-74	Sequence 74, Appl
25	29	28.4	19	4	PCT-US93-07545-75	Sequence 75, Appl
26	29	28.4	20	2	US-08-480-190-191	Sequence 191, App
27	29	28.4	20	2	US-08-749-852-55	Sequence 55, Appl
28	29	28.4	20	2	US-08-488-379-191	Sequence 191, App
29	29	28.4	20	4	PCT-US93-07545-191	Sequence 191, App
30	28	27.5	15	1	US-08-625-691-1	Sequence 1, Appli
31	28	27.5	17	2	US-08-480-190-193	Sequence 193, App
32	28	27.5	17	2	US-08-488-379-193	Sequence 193, App
33	28	27.5	17	4	PCT-US93-07545-193	Sequence 193, App
34	27	26.5	13	4	PCT-US93-06751-102	Sequence 102, App
35	27	26.5	19	3	US-08-792-832A-40	Sequence 40, Appl
36	26	25.5	16	2	US-08-934-222-60	Sequence 60, Appl
37	26	25.5	16	2	US-08-933-402-60	Sequence 60, Appl
38	26	25.5	16	2	US-09-207-621-60	Sequence 60, Appl
39	26	25.5	16	2	US-08-532-818-60	Sequence 60, Appl
40	26	25.5	16	3	US-09-231-797-60	Sequence 60, Appl
41	26	25.5	16	3	US-08-934-224-60	Sequence 60, Appl
42	26	25.5	16	3	US-08-933-843-60	Sequence 60, Appl
43	26	25.5	19	3	US-08-792-832A-34	Sequence 34, Appl
44	25.5	25.0	20	1	US-08-218-025A-8	Sequence 8, Appli
45	25	24.5	17	1	US-07-894-212A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT   1
US-08-472-228A-21
; Sequence 21, Application US/08472228A
; Patent No. 5807978
;  GENERAL INFORMATION:
;    APPLICANT:  Kokolus, William J.
;    APPLICANT:  Fritsche, Herbert A.
;    APPLICANT:  Johnston, Dennis A.
;    TITLE OF INVENTION:  IMMUNOGENIC PEPTIDES OF PROSTATE
;    TITLE OF INVENTION:  SPECIFIC ANTIGEN
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  United States of America
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/472,228A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  UTSC:388
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;      TELEX:  79-0924
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-472-228A-21


  Query Match             29.4%;  Score 30;  DB 1;  Length 19;
  Best Local Similarity   75.0%;  Pred. No. 42;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 RAVMLLGR 13
            |:|:||||
```

```
  Query Match             22.9%;  Score 25;  DB 4;  Length 18;
  Best Local Similarity   46.2%;  Pred. No. 1.8e+03;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 EREPGWLKQLFTN 20
           |:||  | |   |
Db       3 EQEPFALXQFIIN 15


RESULT  12
Q44850
ID   Q44850      PRELIMINARY;      PRT;    19 AA.
AC   Q44850;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   HYPOTHETICAL PROTEIN (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RA   Akins D.R., Popova T., Brusca J., Goldberg M.L., Li M., Baker S.C.,
RA   Norgard M.V., Radolf J.D.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L31425; AAA64901.1; -.
KW   Hypothetical protein.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2321 MW;  C821BF312DB41512 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLHLQGEREP 11
           |:|||    |
Db       4 LIHLQSILHP 13


RESULT  13
Q9UJI8
ID   Q9UJI8      PRELIMINARY;      PRT;    19 AA.
AC   Q9UJI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ333A15.2 (PROSTAGLANDIN E RECEPTOR (EP3E)) (FRAGMENT).
GN   PTGER3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Frankland J.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031429; CAB52457.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2245 MW;  142450379B298FEC CRC64;


  Query Match             22.9%;  Score 25;  DB 4;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LHLQGEREP 11
           ||: |: :|
Db       8 LHIVGKYKP 16


RESULT  14
Q9ZYW2
ID   Q9ZYW2      PRELIMINARY;      PRT;    19 AA.
AC   Q9ZYW2;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME OXIDASE II (FRAGMENT).
OS   Dolopsidea sp.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC   Ichneumonoidea; Braconidae; Rhyssalinae; Dolopsidea.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 99152621.
RA   Dowton M., Austin A.D.;
RT   "Evolutionary dynamics of a mitochondrial rearrangement 'hot spot' in
RT   the Hymenoptera.";
RL   Mol. Biol. Evol. 16:298-309(1999).
DR   EMBL; AF034599; AAC79747.1; -.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2475 MW;  77E78DECE3D4399A CRC64;


  Query Match             22.9%;  Score 25;  DB 8;  Length 19;
  Best Local Similarity   66.7%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      13 WLKQLF 18
           |||: |
Db      14 WLKKWF 19


RESULT  15
O77893
ID   O77893      PRELIMINARY;      PRT;    11 AA.
AC   O77893;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MHC CLASS II B LOCUS 10 (FRAGMENT).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC   Percomorpha; Perciformes; Labroidei; Cichlidae; Oreochromis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98315113.
RA   Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA   Sueltmann H., Figueroa F., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT   class II B loci.";
RL   Genetics 149:1527-1547(1998).
DR   EMBL; AF050003; AAC41342.1; -.
KW   MHC.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1296 MW;  68775B73786B572B CRC64;


  Query Match             22.0%;  Score 24;  DB 7;  Length 11;
  Best Local Similarity   33.3%;  Pred. No. 1.6e+03;
  Matches    2;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 WLKQLF 18
           |:::|:
Db       1 WIRELY 6


Search completed: December 21, 2000, 08:35:37
Job time: 600 sec
```

```
RESULT   7
Q9T2R9
ID   Q9T2R9      PRELIMINARY;      PRT;    17 AA.
AC   Q9T2R9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOCHROME-C REDUCTASE 55 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS   Solanum tuberosum (Potato).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RA   Braun H.P., Kruft V., Schmitz U.K.;
RL   Planta 193:99-106(1994).
SQ   SEQUENCE   17 AA;  1758 MW;  F95F7BF7940F5F21 CRC64;


  Query Match             23.9%;  Score 26;  DB 8;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 LLLHLQG 7
          |:||: |
Db      4 LMLHIDG 10


RESULT   8
Q9UCG3
ID   Q9UCG3      PRELIMINARY;      PRT;    19 AA.
AC   Q9UCG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ALPHA 2-PLASMIN INHIBITOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93249387.
RA   Bangert K., Johnsen A.H., Christensen U., Thorsen S.;
RT   "Different N-terminal forms of alpha 2-plasmin inhibitor in human
RT   plasma.";
RL   Biochem. J. 291:623-625(1993).
SQ   SEQUENCE   19 AA;  2065 MW;  41352BF04D1EEAE9 CRC64;


  Query Match             23.4%;  Score 25.5;  DB 4;  Length 19;
  Best Local Similarity   63.6%;  Pred. No. 1.6e+03;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      2 LLHLQGEREPG 12
          || | | :|||
Db     10 LLKL-GNQEPG 19


RESULT   9
Q16045
ID   Q16045      PRELIMINARY;      PRT;    14 AA.
AC   Q16045;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   D3 DOPAMINE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93326145.
RA   Nagai Y., Ueno S., Saeki Y., Soga F., Yanagihara T.;
RT   "Expression of the D3 dopamine receptor gene and a novel variant
RT   transcript generated by alternative splicing in human peripheral blood
RT   lymphocytes.";
RL   Biochem. Biophys. Res. Commun. 194:368-374(1993).
DR   EMBL; S63845; CAB32270.1; -.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1586 MW;  EA310BEFE94CF1B1 CRC64;


  Query Match             22.9%;  Score 25;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LLHLQ 6
          |||||
Db      9 LLHLQ 13


RESULT  10
Q9R5F6
ID   Q9R5F6      PRELIMINARY;      PRT;    18 AA.
AC   Q9R5F6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   UREASE SMALL SUBUNIT (FRAGMENT).
OS   Helicobacter mustelae.
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93084378.
RA   Turbett G.R., Hoj P.B., Horne R., Mee B.J.;
RT   "Purification and characterization of the urease enzymes of
RT   Helicobacter species from humans and animals.";
RL   Infect. Immun. 60:5259-5266(1992).
SQ   SEQUENCE   18 AA;  2060 MW;  29C8E0AB77E21805 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGE 8
          ::||  ||
Db      9 MMLHYAGE 16


RESULT  11
Q9UC87
ID   Q9UC87      PRELIMINARY;      PRT;    18 AA.
AC   Q9UC87;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   31 KDA EXTRACELLULAR MATRIX-ASSOCIATED SERINE PROTEINASE INHIBITOR
DE   (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95177668.
RA   Rao C.N., Liu Y.Y., Peavey C.L., Woodley D.T.;
RT   "Novel extracellular matrix-associated serine proteinase inhibitors
RT   from human skin fibroblasts.";
RL   Arch. Biochem. Biophys. 317:311-314(1995).
SQ   SEQUENCE   18 AA;  1968 MW;  BA65E727DB150CD6 CRC64;
```

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95337087.
RA   Thulin C.D., Walsh K.A.;
RT   "Identification of the amino terminus of human filaggrin using
RT   differential LC/MS techniques: implications for profilaggrin
RT   processing.";
RL   Biochemistry 34:8687-8692(1995).
SQ   SEQUENCE   20 AA;  2199 MW;  0FD0856B199AB332 CRC64;


  Query Match             26.6%;  Score 29;  DB 4;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 4.8e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 HLQGEREPGW 13
           | | |   ||
Db       5 HEQSESSHGW 14


RESULT   3
Q9PRJ4
ID   Q9PRJ4      PRELIMINARY;      PRT;     9 AA.
AC   Q9PRJ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BRADYKININ.
OS   Lepisosteus osseus (Long-nosed gar), and Amia calva (Bowfin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC   Lepisosteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95380361.
RA   Conlon J.M., Platzack B., Marra L.E., Youson J.H., Olson K.R.;
RT   "Isolation and biological activity of [Trp5]bradykinin from the plasma
RT   of the phylogenetically ancient fish, the bowfin and the longnosed
RT   gar.";
RL   Peptides 16:485-489(1995).
SQ   SEQUENCE   9 AA;  1099 MW;  3393D775A3786777 CRC64;


  Query Match             25.7%;  Score 28;  DB 13;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 REPGW 13
           | |||
Db       1 RPPGW 5


RESULT   4
O18502
ID   O18502      PRELIMINARY;      PRT;    14 AA.
AC   O18502;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   LYSOPHOSPHOLIPASE HOMOLOG (FRAGMENT).
GN   SMLPLH.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGYPTIAN;
RA   Hamdan F.F., Ribeiro P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF006679; AAC62255.1; -.
FT   NON_TER      14      14
SQ   SEQUENCE   14 AA;  1541 MW;  48B847C2E5D89177 CRC64;


  Query Match             25.7%;  Score 28;  DB 5;  Length 14;
  Best Local Similarity   38.5%;  Pred. No. 4.7e+02;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGEREPGW 13
           : ||  |:   ||
Db       2 IFLHGLGDTGHGW 14


RESULT   5
Q9R4W5
ID   Q9R4W5      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4W5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   30 KDA MAJOR HEAT SHOCK PROTEIN (FRAGMENT).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95020803.
RA   Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,
RA   Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
RT   "Heat shock protein produced by Helicobacter pylori.";
RL   Microbiol. Immunol. 38:403-405(1994).
DR   INTERPRO; IPR002026; -.
DR   PFAM; PF00547; urease_gamma; 1.
SQ   SEQUENCE   20 AA;  2302 MW;  29C9DFBFD6D21805 CRC64;


  Query Match             24.8%;  Score 27;  DB 2;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 9.8e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGE 8
           |:||  ||
Db      11 LMLHYAGE 18


RESULT   6
Q9X3I3
ID   Q9X3I3      PRELIMINARY;      PRT;    17 AA.
AC   Q9X3I3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B (FRAGMENT).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070176; AAD20791.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   17 AA;  1953 MW;  630843039ADD51B4 CRC64;


  Query Match             23.9%;  Score 26;  DB 2;  Length 17;
  Best Local Similarity   33.3%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGEREPG 12
           :|:|   :|: |
Db       1 MLMHFLXDRKQG 12
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:36 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-6
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWLKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	26.6	10	13	Q9PRZ1	Q9prz1 oncorhynchu
2	29	26.6	20	4	Q9UC71	Q9uc71 homo sapien
3	28	25.7	9	13	Q9PRJ4	Q9prj4 lepisosteus
4	28	25.7	14	5	O18502	O18502 schistosoma
5	27	24.8	20	2	Q9R4W5	Q9r4w5 helicobacte
6	26	23.9	17	2	Q9X3I3	Q9x3i3 prochloroco
7	26	23.9	17	8	Q9T2R9	Q9t2r9 solanum tub
8	25.5	23.4	19	4	Q9UCG3	Q9ucg3 homo sapien
9	25	22.9	14	4	Q16045	Q16045 homo sapien
10	25	22.9	18	2	Q9R5F6	Q9r5f6 helicobacte
11	25	22.9	18	4	Q9UC87	Q9uc87 homo sapien
12	25	22.9	19	2	Q44850	Q44850 borrelia bu
13	25	22.9	19	4	Q9UJI8	Q9uji8 homo sapien
14	25	22.9	19	8	Q9ZYW2	Q9zyw2 dolopsidea
15	24	22.0	11	7	O77893	O77893 oreochromis
16	24	22.0	12	12	O12088	O12088 caprine art
17	24	22.0	12	12	O12074	O12074 caprine art
18	24	22.0	12	12	O12076	O12076 caprine art
19	24	22.0	12	12	O12078	O12078 caprine art
20	24	22.0	12	12	O12080	O12080 caprine art
21	24	22.0	12	12	O12082	O12082 caprine art
22	24	22.0	12	12	O12084	O12084 caprine art
23	24	22.0	12	12	O12086	O12086 caprine art
24	24	22.0	12	12	O12090	O12090 caprine art
25	24	22.0	12	12	O12092	O12092 caprine art
26	24	22.0	12	12	O12094	O12094 caprine art
27	24	22.0	12	12	O12106	O12106 caprine art
28	24	22.0	12	12	O12108	O12108 caprine art
29	24	22.0	12	12	O12110	O12110 caprine art
30	24	22.0	12	12	O12112	O12112 caprine art
31	24	22.0	12	12	O12114	O12114 caprine art
32	24	22.0	12	12	O12116	O12116 caprine art
33	24	22.0	12	12	O12118	O12118 caprine art
34	24	22.0	15	2	Q46963	Q46963 escherichia
35	24	22.0	15	2	Q9R4U7	Q9r4u7 acinetobact
36	24	22.0	16	4	Q9UCK9	Q9uck9 homo sapien
37	24	22.0	16	4	Q9UCH1	Q9uch1 homo sapien
38	24	22.0	19	2	Q47079	Q47079 escherichia
39	24	22.0	19	10	Q9S8W5	Q9s8w5 avena sativ
40	24	22.0	20	10	Q9S930	Q9s930 glycine max
41	23	21.1	11	5	Q9V7K6	Q9v7k6 drosophila
42	23	21.1	13	4	Q9Y674	Q9y674 homo sapien
43	23	21.1	18	13	Q9PRM7	Q9prm7 xenopus. ac
44	23	21.1	19	10	Q40183	Q40183 lemna gibba
45	22.5	20.6	20	2	Q9R4A6	Q9r4a6 hydrogenoba

ALIGNMENTS

RESULT 1
Q9PRZ1
ID Q9PRZ1 PRELIMINARY; PRT; 10 AA.
AC Q9PRZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE LYSYL-BRADYKININ=VASOACTIVE PEPTIDE KALLIDIN HOMOLOG.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 94039817.
RA Conlon J.M., Olson K.R.;
RT "Purification of a vasoactive peptide related to lysyl-bradykinin from
RT trout plasma.";
RL FEBS Lett. 334:75-78(1993).
SQ SEQUENCE 10 AA; 1193 MW; 33C59075A3786777 CRC64;

Query Match 26.6%; Score 29; DB 13; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 8 EREPGW 13
:| |||
Db 1 KRPPGW 6

RESULT 2
Q9UC71
ID Q9UC71 PRELIMINARY; PRT; 20 AA.
AC Q9UC71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE FILAGGRIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Search completed: December 21, 2000, 08:32:56
Job time: 439 sec

```
RA   Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA   Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT   "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT   dogfish brain provides insight into GnRH evolution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
CC   -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC   -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR   INTERPRO; IPR002012; -.
DR   PFAM; PF00446; GnRH; 1.
DR   PROSITE; PS00473; GNRH; 1.
KW   Hormone; Amidation; Hypothalamus.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1204 MW;  284B32337871F5A3 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 9.8e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     12 GWL 14
          |||
Db      6 GWL 8


RESULT  13
Q2OG_COMTE
ID   Q2OG_COMTE     STANDARD;      PRT;    10 AA.
AC   P80466;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   QUINOLINE 2-OXIDOREDUCTASE, GAMMA CHAIN (EC 1.-.-.-) (FRAGMENT).
OS   Comamonas testosteroni (Pseudomonas testosteroni).
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=63;
RX   MEDLINE; 96035889.
RA   Schach S., Tshisuaka B., Fetzner S., Lingens F.;
RT   "Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT   dioxygenase from Comamonas testosteroni 63. The first two enzymes in
RT   quinoline and 3-methylquinoline degradation.";
RL   Eur. J. Biochem. 232:536-544(1995).
CC   -!- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)2-OXO-
CC       1,2-DIHYDROQUINOLINE.
CC   -!- COFACTOR: FAD AND MOLYBDENUM.
CC   -!- PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
CC       (3-METHYL-)QUINOLINE.
CC   -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC       TWO GAMMA CHAINS (PROBABLE).
KW   Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT   NON_TER       10     10
SQ   SEQUENCE   10 AA;  1153 MW;  C848CE64433B1DC6 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   42.9%;  Pred. No. 9.8e+02;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 LQGEREP 11
          :| |: |
Db      2 IQAEKNP 8


RESULT  14
CXET_CONTE
ID   CXET_CONTE     STANDARD;      PRT;    13 AA.
AC   P81755;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   EPSILON-CONOTOXIN TXIX.
OS   Conus textile (Cloth-of-gold cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
RN   [1]
RP   SEQUENCE, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY NMR.
RC   TISSUE=VENOM;
RX   MEDLINE; 99254114.
RA   Rigby A.C., Lucas-Meunier E., Kalume D.E., Czerwiec E., Hambe B.,
RA   Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,
RA   Furie B.C., Furie B., Stenflo J.P.;
RT   "A conotoxin from Conus textile with unusual posttranslational
RT   modifications reduces presynaptic Ca2+ influx.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).
CC   -!- FUNCTION: CONOTOXIN WHICH ACTS AT PRESYNAPTIC MEMBRANES, BLOCKING
CC       THE CALCIUM CHANNELS.
CC   -!- PTM: O-GLYCAN CONSISTS OF THE DISACCHARIDE GAL-GALNAC.
DR   PDB; 1WCT; 08-JUN-99.
KW   Presynaptic neurotoxin; Calcium channel inhibitor; Venom; Vitamin K;
KW   Gamma-carboxyglutamic acid; Glycoprotein; Hydroxylation; Bromination;
KW   3D-structure.
FT   DISULFID      2      8
FT   DISULFID      3      9
FT   MOD_RES       1      1       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       4      4       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       7      7       BROMINATION.
FT   MOD_RES      13     13       HYDROXYLATION.
FT   CARBOHYD     10     10       O-LINKED (GALNAC. . .).
SQ   SEQUENCE   13 AA;  1388 MW;  386C9E1C74AFA378 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     10 EPGW 13
          | ||
Db      4 EDGW 7


RESULT  15
FARB_ASCSU
ID   FARB_ASCSU     STANDARD;      PRT;    13 AA.
AC   P43173;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE AF11.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95380362.
RA   Cowden C., Stretton A.O.W.;
RT   "Eight novel FMRFamide-like neuropeptides isolated from the nematode
RT   Ascaris suum.";
RL   Peptides 16:491-500(1995).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES      13     13       AMIDATION.
SQ   SEQUENCE   13 AA;  1495 MW;  9CAEC650D6886B05 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      7 GEREPGWLK 15
          |  || :|:
Db      4 GISEPNFLR 12
```

```
CC   -!- FUNCTION: SMOOTH MUSCLE CONTRACTION.
KW   Bradykinin; Vasodilator.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1573 MW;  2673CB3D83ECC867 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 9e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 QGEREPGW 13
            | :| ||:
Db       1 QZKRPPGF 8


RESULT   9
HPA1_RANES
ID   HPA1_RANES     STANDARD;      PRT;    13 AA.
AC   P32415;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   HEMOLYTIC PROTEIN A1 (FRAGMENT).
OS   Rana esculenta (Edible frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN SECRETION;
RX   MEDLINE; 90198965.
RA   Simmaco M., de Biase D., Severini C., Aita M., Erspamer G.F.,
RA   Barra D., Bossa F.;
RT   "Purification and characterization of bioactive peptides from skin
RT   extracts of Rana esculenta.";
RL   Biochim. Biophys. Acta 1033:318-323(1990).
CC   -!- FUNCTION: SHOWS HEMOLYTIC ACTIVITY.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: SKIN.
DR   PIR; S09018; S09018.
KW   Amphibian skin; Amidation; Hemolysis.
FT   MOD_RES      13     13       AMIDATION.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1390 MW;  C6BA768B9DFE587D CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   71.4%;  Pred. No. 9e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           | | |||
Db       7 GILSQLF 13


RESULT  10
NEUT_CHICK
ID   NEUT_CHICK     STANDARD;      PRT;    13 AA.
AC   P13724;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   NEUROTENSIN (NT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88063566.
RA   Iwabuchi H., Komori S., Ohashi H., Kimura S.;
RT   "The amino acid sequence of a smooth muscle-contracting peptide from
RT   chicken rectum. Identity to chicken neurotensin.";
RL   Jpn. J. Pharmacol. 44:455-459(1987).
CC   -!- FUNCTION: SMOOTH MUSCLE-CONTRACTING PEPTIDE.
DR   PIR; A28505; A28505.
KW   Vasoactive.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1608 MW;  4C949E714C410DD3 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 9e+02;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LHLQGEREP 11
            ||:   | |
Db       2 LHVNKARRP 10


RESULT  11
GON3_ONCKE
ID   GON3_ONCKE     STANDARD;      PRT;    10 AA.
AC   P20367;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GONADOLIBERIN III (GONADOTROPIN-RELEASING HORMONE III) (GNRH-III) (LH-
DE   RH III) (LULIBERIN III).
OS   Oncorhynchus keta (Chum salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83195140.
RA   Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
RT   "Characterization of a teleost gonadotropin-releasing hormone.";
RL   Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
CC   -!- FUNCTION: STIMULATES THE SECRETION OF BOTH LUTEINIZING AND
CC       FOLLICLE-STIMULATING HORMONES.
CC   -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR   PIR; A21114; A21114.
DR   INTERPRO; IPR002012; -.
DR   PFAM; PF00446; GnRH; 1.
DR   PROSITE; PS00473; GNRH; 1.
KW   Hormone; Amidation; Hypothalamus.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1230 MW;  284B3233786B45A3 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 9.8e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 GWL 14
           |||
Db       6 GWL 8


RESULT  12
GONL_SQUAC
ID   GONL_SQUAC     STANDARD;      PRT;    10 AA.
AC   P27429;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   GONADOLIBERIN (GONADOTROPIN-RELEASING HORMONE) (GNRH) (LH-RH)
DE   (LULIBERIN).
OS   Squalus acanthias (Spiny dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Squalea; Squaloidei; Squalidae; Squalus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 92335300.
```

```
DE   (DCC II).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 90046758.
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; B33995; B33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1169 MW;  916036786771A9D1 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   6
AROQ_AMYME
ID   AROQ_AMYME      STANDARD;      PRT;     20 AA.
AC   P46380;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10) (3-DEHYDROQUINASE)
DE   (TYPE II DHQASE) (FRAGMENT).
GN   AROQ.
OS   Amycolatopsis methanolica.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC   Amycolatopsis.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=NCIB 11946;
RX   MEDLINE; 93123995.
RA   Euverink G.J.W., Hessels G.I., Vrijbloed J.W., Coggins J.R.,
RA   Dijkhuizen L.;
RT   "Purification and characterization of a dual function
RT   3-dehydroquinate dehydratase from Amycolatopsis methanolica.";
RL   J. Gen. Microbiol. 138:2449-2457(1992).
CC   -!- FUNCTION: CATALYZE A TRANS-DEHYDRATION VIA AN ENOLATE
CC       INTERMEDIATE. IS INVOLVED IN BOTH THE CATABOLISM OF QUINATE AND
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS. HAS A TEMPERATURE
CC       OPTIMAL OF 76 DEGREES CELSIUS AND PH OPTIMAL OF 9.0. TYPE II
CC       ENZYMES ARE THERMOSTABLE.
CC   -!- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE = 3-DEHYDROSHIKIMATE + H(2)O.
CC   -!- PATHWAY: THIRD STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: HOMODODECAMER.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-II 3-DEHYDROQUINASE FAMILY.
DR   INTERPRO; IPR001874; -.
DR   PROSITE; PS01029; DEHYDROQUINASE_II; PARTIAL.
KW   Quinate metabolism; Aromatic amino acid biosynthesis; Lyase.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2197 MW;  C24AA183E5CFDF0A CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GEREP 11
            |:|||
Db      15 GKREP 19


RESULT   7
LPK1_LOCMI
ID   LPK1_LOCMI      STANDARD;      PRT;     16 AA.
AC   P20404;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   LOCUSTAPYROKININ 1 (LOM-PK-1).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 91224474.
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, primary structure, and synthesis of locustapyrokinin: a
RT   myotropic peptide of Locusta migratoria.";
RL   Gen. Comp. Endocrinol. 81:97-104(1991).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   PIR; A49761; A49761.
DR   INTERPRO; IPR001484; -.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1827 MW;  A7178BBDCA0AFDD6 CRC64;


  Query Match             21.1%;  Score 23;  DB 1;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 7.9e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           || :| |
Db       6 GWPQQPF 12


RESULT   8
BRK_PARID
ID   BRK_PARID      STANDARD;      PRT;     13 AA.
AC   P42717;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   WASPKININ.
OS   Parapolybia indica.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Polistinae; Parapolybia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Toki T., Yasuhara T., Nakajima T.;
RT   "Isolation and sequential analysis of peptides on the venom sac of
RT   Parapolybia indica.";
RL   Eisei Dobutsu 39:105-111(1988).
```

```
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  86786771A9D1A736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   3
RPCH_PANBO
ID   RPCH_PANBO     STANDARD;      PRT;     8 AA.
AC   P08939;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RED PIGMENT CONCENTRATING HORMONE (RPCH).
OS   Pandalus borealis (Northern red shrimp).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidea;
OC   Pandalidae; Pandalus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 75054965.
RA   Fernlund P.;
RT   "Structure of the red-pigment-concentrating hormone of the shrimp,
RT   Pandalus borealis.";
RL   Biochim. Biophys. Acta 371:304-311(1974).
CC   -!- FUNCTION: THIS HORMONE ADAPTS THE ANIMAL TO LIGHT BACKGROUNDS BY
CC       STIMULATING CONCENTRATION OF THE PIGMENT OF ITS RED BODY-
CC       CHROMATOPHORES.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S07139; S07139.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Pigment; Hormone; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  948 MW;  86786775B9C44736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   4
HTF_NAUCI
ID   HTF_NAUCI      STANDARD;      PRT;     10 AA.
AC   P10939;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPERTREHALOSAEMIC HORMONE (HTH) (HYPERTREHALOSAEMIC NEUROPEPTIDE).
OS   Nauphoeta cinerea (Cinereous cockroach) (Gray cockroach),
OS   Leucophaea maderae (Madeira cockroach),
OS   Blattella germanica (German cockroach), and
OS   Gromphadorina portentosa (Cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Nauphoeta.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=N.CINEREA; TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 87100208.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Amino acid sequence of a hypertrehalosaemic neuropeptide from the
RT   corpus cardiacum of the cockroach, Nauphoeta cinerea.";
RL   Biochem. Biophys. Res. Commun. 141:774-781(1986).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=L.MADERAE, G.PORTENTOSA, AND B.GERMANICA;
RX   MEDLINE; 90253659.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=B.GERMANICA;
RX   MEDLINE; 91179584.
RA   Veenstra J.A., Camps F.;
RT   "Structure of the hypertrehalosemic neuropeptide of the German
RT   cockroach, Blattella germanica.";
RL   Neuropeptides 15:107-109(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A26381; A26381.
DR   PIR; S08997; S08997.
DR   PIR; S08998; S08998.
DR   PIR; S09137; S09137.
DR   PIR; A60421; A60421.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1092 MW;  056236786775B9C4 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   5
HTF_TABAT
ID   HTF_TABAT      STANDARD;      PRT;     10 AA.
AC   P14596;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   HYPERTREHALOSAEMIC FACTOR (HOTH) (DIPTERAN CORPORA CARDIACA FACTOR II)
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:55 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-6
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWLKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.7	19	1	TRP3_LEUMA	P81735 leucophaea
2	24	22.0	8	1	AKH_TABAT	P14595 tabanus atr
3	24	22.0	8	1	RPCH_PANBO	P08939 pandalus bo
4	24	22.0	10	1	HTF_NAUCI	P10939 nauphoeta c
5	24	22.0	10	1	HTF_TABAT	P14596 tabanus atr
6	24	22.0	20	1	AROQ_AMYME	P46380 amycolatops
7	23	21.1	16	1	LPK1_LOCMI	P20404 locusta mig
8	22	20.2	13	1	BRK_PARID	P42717 parapolybia
9	22	20.2	13	1	HPA1_RANES	P32415 rana escule
10	22	20.2	13	1	NEUT_CHICK	P13724 gallus gall
11	21	19.3	10	1	GON3_ONCKE	P20367 oncorhynchu
12	21	19.3	10	1	GONL_SQUAC	P27429 squalus aca
13	21	19.3	10	1	Q2OG_COMTE	P80466 comamonas t
14	21	19.3	13	1	CXET_CONTE	P81755 conus texti
15	21	19.3	13	1	FARB_ASCSU	P43173 ascaris suu
16	21	19.3	15	1	VORA_METTM	P80907 methanobact
17	21	19.3	17	1	CHH3_BOMMO	P20729 bombyx mori
18	21	19.3	20	1	CRTC_SPIOL	P30806 spinacia ol
19	20	18.3	12	1	PORD_METTM	P80903 methanobact
20	20	18.3	15	1	TA1_TREBR	P34070 tremella br
21	19.5	17.9	20	1	COG1_PARCM	P20731 paralithode
22	19	17.4	7	1	UH11_RAT	P56576 rattus norv
23	19	17.4	8	1	CCKN_MACEU	P30369 macropus eu
24	19	17.4	10	1	CAER_LITXA	P56264 litoria xan
25	19	17.4	10	1	GON1_PETMA	P04378 petromyzon
26	19	17.4	12	1	RS19_ELYEP	Q47881 elm yellows
27	19	17.4	13	1	NEUT_RANTE	P41536 rana tempor
28	19	17.4	14	1	RS19_CLOPP	Q46228 clover prol
29	19	17.4	14	1	RS19_LOWBP	Q48878 loofah witc
30	19	17.4	17	1	GAST_MACMU	P33714 macaca mula
31	19	17.4	17	1	LCK_RAT	Q01621 rattus norv
32	19	17.4	18	1	NPA_BOVIN	P15506 bos taurus
33	19	17.4	19	1	FIBB_HORSE	P14471 equus cabal
34	19	17.4	19	1	PYRB_PSEFL	P56585 pseudomonas
35	18	16.5	8	1	AKH_LIBAU	P25418 libellula a
36	18	16.5	8	1	HTF1_PERAM	P04548 periplaneta
37	18	16.5	8	1	HTF2_PERAM	P04549 periplaneta
38	18	16.5	8	1	HTF_TENMO	P25419 tenebrio mo
39	18	16.5	8	1	UF06_MOUSE	P38644 mus musculu
40	18	16.5	10	1	HTF1_ROMMI	P18110 romalea mic
41	18	16.5	10	1	HTF2_CARMO	P11385 carausius m
42	18	16.5	11	1	BRK_MEGFL	P12797 megascolia
43	18	16.5	11	1	TEMK_RANTE	P56922 rana tempor
44	18	16.5	11	1	TEML_RANTE	P56923 rana tempor
45	18	16.5	13	1	TP13_PHYRO	P04096 phyllomedus

ALIGNMENTS

RESULT 1
TRP3_LEUMA
ID TRP3_LEUMA STANDARD; PRT; 19 AA.
AC P81735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TACHYKININ-RELATED PEPTIDE 3 (LEMTRP 3).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
RN [1]
RP SEQUENCE.
RC TISSUE=MIDGUT;
RX MEDLINE; 97053012.
RA Muren J.E., Naessel D.R.;
RT "Isolation of five tachykinin-related peptides from the midgut of
RT the cockroach Leucophaea madera: existence of N-terminally extended
RT isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 19 19 AMIDATION.
SQ SEQUENCE 19 AA; 1930 MW; 99B5471A011625E5 CRC64;

Query Match 25.7%; Score 28; DB 1; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      7 GEREPG 12
          ||| ||
Db      2 GERAPG 7
```

RESULT 2
AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ADIPOKINETIC HORMONE (AKH) (DIPTERAN CORPORA CARDIACA FACTOR I)
DE (DCC I).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC Tabanidae; Tabanus.
RN [1]
RP SEQUENCE.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE; 90046758.

A;Accession: A58620
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 14
A43405
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC 3.
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Mar-2000
C;Accession: A43405
R;Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilkis, S.J.; Bartrons, R.
J. Biol. Chem. 267, 17939-17943, 1992
A;Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence for
A;Reference number: A43405; MUID:92388154
A;Accession: A43405
A;Molecule type: protein
A;Residues: 1-10 <VEN>
C;Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phosp
C;Keywords: phosphoric monoester hydrolase; phosphotransferase

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 6.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLQGE 8
||:||
Db 5 HLKGE 9

RESULT 15
A60421
hypertrehalosemic hormone - German cockroach
N;Alternate names: Bld-HrTH
C;Species: Blattella germanica (German cockroach)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 31-Oct-1997
C;Accession: A60421; S09137
R;Veenstra, J.A.; Camps, F.
Neuropeptides 15, 107-109, 1990
A;Title: Structure of the hypertrehalosemic neuropeptide of the German cockroach, Blatte
A;Reference number: A60421; MUID:91179584
A;Accession: A60421
A;Molecule type: protein
A;Residues: 1-10 <VEE>
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659
A;Accession: S09137
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

Search completed: December 21, 2000, 08:30:05
Job time: 269 sec

```
C;Function:
A;Description: involved in turnover of connective tissues
C;Keywords: serine proteinase inhibitor


  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   46.2%;  Pred. No. 8.5e+02;
  Matches    6;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 EREPGWLKQLFTN 20
           |:||  | |   |
Db       3 EQEPFALXQFIIN 15


RESULT   9
A34626
RPCH-related neuropeptide - ferruginous spindle
C;Species: Fusinus ferrugineus (ferruginous spindle)
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C;Accession: A34626
R;Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake, Y.;
Biochem. Biophys. Res. Commun. 167, 273-279, 1990
A;Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
A;Reference number: A34626; MUID:90179762
A;Accession: A34626
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <KUR>
C;Keywords: neuropeptide


  Query Match             22.0%;  Score 24;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       2 PGW 4


RESULT  10
A61348
red pigment-concentrating hormone - northern shrimp
N;Alternate names: blanching hormone
C;Species: Pandalus borealis (northern shrimp)
C;Date: 02-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 07-May-1999
C;Accession: A61348; S07139
R;Fernlund, P.; Josefsson, L.
Science 177, 173-175, 1972
A;Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.
A;Reference number: A61348; MUID:72228738
A;Accession: A61348
A;Molecule type: protein
A;Residues: 1-8 <FER1>
R;Fernlund, P.
Biochim. Biophys. Acta 371, 304-311, 1974
A;Title: Structure of the red-pigment-concentrating hormone of the shrimp, Pandalus bore
A;Reference number: S07139; MUID:75054965
A;Accession: S07139
A;Molecule type: protein
A;Residues: 'E',2-8 <FER2>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C;Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pi
zed pigment-containing cells.
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutami
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  11
A33995
adipokinetic hormone - black horse fly
C;Species: Tabanus atratus (black horse fly)
C;Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 31-Oct-1997
C;Accession: A33995
R;Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Nachman, R.J.; Vogel, V.W.; Zhan
Proc. Natl. Acad. Sci. U.S.A. 86, 8161-8164, 1989
A;Title: Primary structure of two neuropeptide hormones with adipokinetic and hypotre
A;Reference number: A33995; MUID:90046758
A;Accession: A33995
A;Molecule type: protein
A;Residues: 1-8 <JAF>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyrogluta
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F;8/Modified site: amidated carboxyl end (Trp) #status predicted


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  12
S55310
adipokinetic hormone - damselfly (Pseudagrion inconspicuum)
N;Alternate names: Psi-AKH
C;Species: Pseudagrion inconspicuum
C;Date: 19-Mar-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: S55310
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
A;Accession: S55310
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental


  Query Match             22.0%;  Score 24;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT  13
A58620
adipokinetic hormone - damselfly (Ischnura senegalensis)
C;Species: Ischnura senegalensis
C;Date: 28-Oct-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: A58620
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
```

G02018
proteasome chain LMP7 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 22-Jun-1999
C;Accession: G02018
R;Kim, T.
submitted to the EMBL Data Library, July 1995
A;Reference number: G09054
A;Accession: G02018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <KIM>
A;Cross-references: EMBL:U32862; NID:g1045468; PIDN:AAA80234.1; PID:g1045469
C;Genetics:
A;Gene: LMP7
C;Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 26.6%; Score 29; DB 2; Length 18;
Best Local Similarity 57.1%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 9 REPGWLK 15
:| ||:|
Db 4 KEDGWVK 10

RESULT 4
PH1788
T cell receptor alpha chain V region (clone 2PBL V alpha 24-4) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1788
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1788
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <POR>

Query Match 24.8%; Score 27; DB 2; Length 15;
Best Local Similarity 71.4%; Pred. No. 3.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 HLQGERE 10
|| |||:
Db 1 HLCGERQ 7

RESULT 5
PH1802
T cell receptor alpha chain V region (clone 3PBL V alpha 24-8) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1802
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>

Query Match 22.9%; Score 25; DB 2; Length 17;
Best Local Similarity 83.3%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 HLQGER 9
|| |||
Db 1 HLCGER 6

RESULT 6
PH1822
T cell receptor alpha chain V region (clone 5PBL V alpha 24-5) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1822
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1822
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>

Query Match 22.9%; Score 25; DB 2; Length 17;
Best Local Similarity 83.3%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 HLQGER 9
|| |||
Db 1 HLCGER 6

RESULT 7
F49215
urease (EC 3.5.1.5) small chain UreA - Helicobacter mustelae (ATCC 43772) (fragment)
C;Species: Helicobacter mustelae
C;Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C;Accession: F49215
R;Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A;Title: Purification and characterization of the urease enzymes of Helicobacter spec
A;Reference number: A49215; MUID:93084378
A;Accession: F49215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <TUR>
A;Note: sequence extracted from NCBI backbone (NCBIP:119487)
C;Superfamily: urease 26K chain; urease 11K chain homology; urease 12K chain homology
C;Keywords: hydrolase

Query Match 22.9%; Score 25; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LLLHLQGE 8
::|| ||
Db 9 MMLHYAGE 16

RESULT 8
S71592
serine proteinase inhibitor, 31K - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 13-Sep-1998
C;Accession: S71592
R;Rao, C.N.; Liu, Y.Y.; Peavey, C.L.; Woodley, D.T.
Arch. Biochem. Biophys. 317, 311-314, 1995
A;Title: Novel extracellular matrix-associated serine proteinase inhibitors from huma
A;Reference number: S71592; MUID:95177668
A;Accession: S71592
A;Molecule type: protein
A;Residues: 1-18 <RAO>

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:04 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-6
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWLKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	27.5	16	2	PH1790	T cell receptor al
2	29	26.6	10	2	S39030	lysyl-bradykinin -
3	29	26.6	18	2	G02018	proteasome chain L
4	27	24.8	15	2	PH1788	T cell receptor al
5	25	22.9	17	2	PH1802	T cell receptor al
6	25	22.9	17	2	PH1822	T cell receptor al
7	25	22.9	18	2	F49215	urease (EC 3.5.1.5
8	25	22.9	18	2	S71592	serine proteinase
9	24	22.0	4	2	A34626	RPCH-related neuro
10	24	22.0	8	2	A61348	red pigment-concen
11	24	22.0	8	2	A33995	adipokinetic hormo
12	24	22.0	8	2	S55310	adipokinetic hormo
13	24	22.0	8	2	A58620	adipokinetic hormo
14	24	22.0	10	2	A43405	6-phosphofructo-2-
15	24	22.0	10	2	A60421	hypertrehalosemic
16	24	22.0	10	2	S08997	hypertrehalosemic
17	24	22.0	10	2	S08998	hypertrehalosemic
18	24	22.0	10	2	A26381	hypertrehalosemic
19	24	22.0	10	2	B33995	hypotrehalosemic h
20	24	22.0	11	2	PD0442	NIPSNAP2 protein -
21	24	22.0	15	2	A26228	spot 42 protein -
22	24	22.0	15	2	PH1329	Ig heavy chain DJ
23	24	22.0	15	2	S38976	chromogranin A - b
24	24	22.0	16	2	PT0282	Ig heavy chain CDR
25	24	22.0	19	2	A41668	integrin alpha-7 c
26	24	22.0	19	2	S29212	protein C - oat (f
27	24	22.0	20	2	A47687	3-dehydroquinate d
28	23	21.1	15	2	S67975	apolipoprotein Cb2
29	23	21.1	16	1	A49761	locustapyrokinin -
30	23	21.1	17	2	PH0778	T-cell receptor al
31	23	21.1	19	2	PH1339	Ig heavy chain DJ
32	22	20.2	7	4	A58725	virotoxin - destro
33	22	20.2	9	2	G56978	collagen alpha 1(I
34	22	20.2	11	2	D56979	collagen alpha 1(I
35	22	20.2	11	2	A33571	follistatin - bovi
36	22	20.2	12	2	PQ0776	NADH dehydrogenase
37	22	20.2	13	2	S09018	hemolytic protein
38	22	20.2	13	2	A28505	neurotensin-like p
39	22	20.2	14	2	PH1347	Ig heavy chain DJ
40	22	20.2	14	2	PH1311	Ig heavy chain DJ
41	22	20.2	14	2	PH1321	Ig heavy chain DJ
42	22	20.2	14	2	PH1305	Ig heavy chain DJ
43	22	20.2	14	2	PH1306	Ig heavy chain DJ
44	22	20.2	15	2	PT0222	Ig heavy chain CDR
45	22	20.2	15	2	PH1314	Ig heavy chain DJ

ALIGNMENTS

RESULT 1
PH1790
T cell receptor alpha chain V region (clone 2PBL V alpha 24-6) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1790
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1790
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16 <POR>

Query Match 27.5%; Score 30; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      4 HLQGEREP 11
          || ||| |
Db      1 HLCGERGP 8
```

RESULT 2
S39030
lysyl-bradykinin - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 19-May-1994 #sequence_revision 19-Apr-1996 #text_change 19-Apr-1996
C;Accession: S39030
R;Conlon, J.M.; Olson, K.R.
FEBS Lett. 334, 75-78, 1993
A;Title: Purification of a vasoactive peptide related to lysyl-bradykinin from trout
A;Reference number: S39030; MUID:94039817
A;Accession: S39030
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <CON>

Query Match 26.6%; Score 29; DB 2; Length 10;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      8 EREPGW 13
          :| |||
Db      1 KRPPGW 6
```

RESULT 3

```
Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  14
PCT-US95-09262-13
; Sequence 13, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-13

  Query Match             32.1%;  Score 35;  DB 4;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  15
US-08-621-803-227
; Sequence 227, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  227:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.388"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-227


  Query Match             30.3%;  Score 33;  DB 2;  Length 10;
  Best Local Similarity   85.7%;  Pred. No. 15;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       1 GWLIQLF 7



Search completed: December 21, 2000, 08:31:46
Job time: 369 sec
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/02465
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noonan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
PCT-US94-02465-40


  Query Match             32.1%;  Score 35;  DB 4;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  12
PCT-US95-00498-40
; Sequence 40, Application PC/TUS9500498
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00498
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00498-40


  Query Match             32.1%;  Score 35;  DB 4;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  13
PCT-US95-00656-40
; Sequence 40, Application PC/TUS9500656
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Materials and Methods
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00656
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00656-40


  Query Match             32.1%;  Score 35;  DB 4;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-13


  Query Match             32.1%;  Score 35;  DB 2;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  10
US-09-119-263-40
; Sequence 40, Application US/09119263
; Patent No. 6054431
;  GENERAL INFORMATION:
;    APPLICANT:  Horwitz, Arnold H.
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/119,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,116
;      FILING DATE:
;      APPLICATION NUMBER:  08/372,783
;      FILING DATE:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-09-119-263-40


  Query Match             32.1%;  Score 35;  DB 3;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT  11
PCT-US94-02465-40
; Sequence 40, Application PC/TUS9402465
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;       TELEFAX:  312-715-1234
;       TELEX:  910-221-5317
;   INFORMATION FOR SEQ ID NO:  40:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "BPI.38"
US-08-473-344-40


  Query Match           32.1%;  Score 35;  DB 1;  Length 14;
  Best Local Similarity  85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT   7
US-08-621-803-39
; Sequence 39, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;     APPLICANT:  Better, Marc D.
;     TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;     TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;     NUMBER OF SEQUENCES:  265
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;       STREET:  6300 Sears Tower, 233 South Wacker Drive
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  United States of America
;       ZIP:  60606-6402
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/621,803
;       FILING DATE:  22-MAR-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25,447
;       REFERENCE/DOCKET NUMBER:  27129/33199
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312/474-6300
;       TELEFAX:  312/474-0448
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  39:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "XMP.38"
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  C-Terminus
;       OTHER INFORMATION:  /label= Amidation
;       OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-39


  Query Match           32.1%;  Score 35;  DB 2;  Length 14;
  Best Local Similarity  85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT   8
US-08-485-445A-40
; Sequence 40, Application US/08485445A
; Patent No. 5856438
;  GENERAL INFORMATION:
;     APPLICANT:  Little, Roger G.
;     TITLE OF INVENTION:  Biologically Active Peptides from
;     TITLE OF INVENTION:  Functional Domains of Bactericidal/
;     TITLE OF INVENTION:  Permeability-Increasing Protein and
;     TITLE OF INVENTION:  Uses Thereof
;     NUMBER OF SEQUENCES:  226
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;       STREET:  Suite 3400, 500 West Madison Street
;       CITY:  Chicago
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60661
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/485,445A
;       FILING DATE:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/209,762
;       FILING DATE:  11-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/183,222
;       FILING DATE:  14-JAN-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McNicholas, Janet M.
;       REGISTRATION NUMBER:  32,918
;       REFERENCE/DOCKET NUMBER:  11018US08/100-224.P4.C1B
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-707-8889
;       TELEFAX:  312-707-9155
;   INFORMATION FOR SEQ ID NO:  40:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  misc_feature
;       OTHER INFORMATION:  "BPI.38"
US-08-485-445A-40


  Query Match           32.1%;  Score 35;  DB 2;  Length 14;
  Best Local Similarity  85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT   9
US-08-621-259A-13
; Sequence 13, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
```

```
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/
;    TITLE OF INVENTION:  Permeability-Increasing Protein and
;    TITLE OF INVENTION:  Uses Thereof
;    NUMBER OF SEQUENCES:  226
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  Suite 3000, 10 S. Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,473A
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-306-473A-40


  Query Match             32.1%;  Score 35;  DB 1;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT   5
US-08-209-762-40
; Sequence 40, Application US/08209762
; Patent No. 5733872
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,762
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5733872nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-209-762-40


  Query Match             32.1%;  Score 35;  DB 1;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 10;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWLKQLF 18
           ||| |||
Db       5 GWLAQLF 11


RESULT   6
US-08-473-344-40
; Sequence 40, Application US/08473344
; Patent No. 5763567
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10  South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,344
;      FILING DATE:  7-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,473
;      FILING DATE:  15-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,762
;      FILING DATE:  11-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-J
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
```

US-08-311-611A-40

Query Match 32.1%; Score 35; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 GWLKQLF 18
||| |||
Db 5 GWLAQLF 11

RESULT 2
US-08-372-783-40
; Sequence 40, Application US/08372783
; Patent No. 5578572
; GENERAL INFORMATION:
; APPLICANT: Horwitz, Arnold H.
; APPLICANT: Lambert, Lewis H.
; APPLICANT: Little, Roger G.
; TITLE OF INVENTION: Anti-Gram-Positive Bacterial Methods and
; TITLE OF INVENTION: Materials
; NUMBER OF SEQUENCES: 237
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,783
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27129/32415
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.38"
US-08-372-783-40

Query Match 32.1%; Score 35; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 GWLKQLF 18
||| |||
Db 5 GWLAQLF 11

RESULT 3
US-08-372-105-40
; Sequence 40, Application US/08372105
; Patent No. 5627153
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.
; APPLICANT: Lim, Edward
; APPLICANT: Lambert, Lewis H.
; APPLICANT: Scannon, Patrick J.
; TITLE OF INVENTION: Anti-Fungal Materials and Methods
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/372,105
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,540
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/209,762
; FILING DATE: 11-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/183,222
; FILING DATE: 14-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 27129/32415
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: "XMP.38"
US-08-372-105-40

Query Match 32.1%; Score 35; DB 1; Length 14;
Best Local Similarity 85.7%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 12 GWLKQLF 18
||| |||
Db 5 GWLAQLF 11

RESULT 4
US-08-306-473A-40
; Sequence 40, Application US/08306473A
; Patent No. 5652332
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:46 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-6
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWLKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	32.1	14	1	US-08-311-611A-40	Sequence 40, Appl
2	35	32.1	14	1	US-08-372-783-40	Sequence 40, Appl
3	35	32.1	14	1	US-08-372-105-40	Sequence 40, Appl
4	35	32.1	14	1	US-08-306-473A-40	Sequence 40, Appl
5	35	32.1	14	1	US-08-209-762-40	Sequence 40, Appl
6	35	32.1	14	1	US-08-473-344-40	Sequence 40, Appl
7	35	32.1	14	2	US-08-621-803-39	Sequence 39, Appl
8	35	32.1	14	2	US-08-485-445A-40	Sequence 40, Appl
9	35	32.1	14	2	US-08-621-259A-13	Sequence 13, Appl
10	35	32.1	14	3	US-09-119-263-40	Sequence 40, Appl
11	35	32.1	14	4	PCT-US94-02465-40	Sequence 40, Appl
12	35	32.1	14	4	PCT-US95-00498-40	Sequence 40, Appl
13	35	32.1	14	4	PCT-US95-00656-40	Sequence 40, Appl
14	35	32.1	14	4	PCT-US95-09262-13	Sequence 13, Appl
15	33	30.3	10	2	US-08-621-803-227	Sequence 227, App
16	33	30.3	10	2	US-08-621-259A-219	Sequence 219, App
17	33	30.3	11	2	US-08-621-803-228	Sequence 228, App
18	33	30.3	11	2	US-08-621-259A-220	Sequence 220, App
19	33	30.3	12	2	US-08-621-803-230	Sequence 230, App
20	33	30.3	12	2	US-08-621-259A-222	Sequence 222, App
21	33	30.3	13	1	US-08-261-660A-33	Sequence 33, Appl
22	33	30.3	13	1	US-08-261-660A-42	Sequence 42, Appl
23	33	30.3	13	4	PCT-US94-06931-33	Sequence 33, Appl
24	33	30.3	13	4	PCT-US94-06931-42	Sequence 42, Appl
25	33	30.3	14	1	US-08-311-611A-15	Sequence 15, Appl
26	33	30.3	14	1	US-08-311-611A-33	Sequence 33, Appl
27	33	30.3	14	1	US-08-311-611A-34	Sequence 34, Appl
28	33	30.3	14	1	US-08-311-611A-35	Sequence 35, Appl
29	33	30.3	14	1	US-08-311-611A-36	Sequence 36, Appl
30	33	30.3	14	1	US-08-311-611A-44	Sequence 44, Appl
31	33	30.3	14	1	US-08-311-611A-45	Sequence 45, Appl
32	33	30.3	14	1	US-08-311-611A-46	Sequence 46, Appl
33	33	30.3	14	1	US-08-311-611A-81	Sequence 81, Appl
34	33	30.3	14	1	US-08-311-611A-89	Sequence 89, Appl
35	33	30.3	14	1	US-08-311-611A-91	Sequence 91, Appl
36	33	30.3	14	1	US-08-311-611A-106	Sequence 106, App
37	33	30.3	14	1	US-08-311-611A-107	Sequence 107, App
38	33	30.3	14	1	US-08-311-611A-109	Sequence 109, App
39	33	30.3	14	1	US-08-311-611A-110	Sequence 110, App
40	33	30.3	14	1	US-08-311-611A-122	Sequence 122, App
41	33	30.3	14	1	US-08-311-611A-142	Sequence 142, App
42	33	30.3	14	1	US-08-372-783-15	Sequence 15, Appl
43	33	30.3	14	1	US-08-372-783-33	Sequence 33, Appl
44	33	30.3	14	1	US-08-372-783-34	Sequence 34, Appl
45	33	30.3	14	4	PCT-US95-09262-62	Sequence 62, Appl

ALIGNMENTS

```
RESULT 1
US-08-311-611A-40
; Sequence 40, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
;   APPLICANT: Cohen, Jonathan
;   APPLICANT: Kung, Ada H.C.
;   APPLICANT: Lambert, Jr., Lewis H.
;   TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
;   TITLE OF INVENTION: Infection by Administration of
;   TITLE OF INVENTION: Bactericidal/Permeability-Increasing
;   TITLE OF INVENTION:
;   NUMBER OF SEQUENCES: 227
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;     STREET: 6300 Sears Tower, 233 South Wacker Drive
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: USA
;     ZIP: 60606-6402
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/311,611A
;     FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/273,401
;     FILING DATE: 11-JUL-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/125,651
;     FILING DATE: 22-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sharp, Jeffrey S.
;     REGISTRATION NUMBER: 31,879
;     REFERENCE/DOCKET NUMBER: 32251
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312/474-6300
;     TELEFAX: 312/474-0448
;     TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 14 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: misc_feature
;     OTHER INFORMATION: "BPI.38"
```

```
Qy      6 LLQMDFGFPKHLL 18
          | : :||   |:|
Db      6 LFKPNFGASLHIL 18
```

Search completed: December 21, 2000, 08:40:02
Job time: 125 sec

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=WHOLE EMBRYOS;
RX   MEDLINE; 98141813.
RA   Peale F.V., Mason K., Hunter A.W., Bothwell M.;
RT   "Multiplex display polymerase chain reaction amplifies and resolves
RT   related sequences sharing a single moderately conserved domain.";
RL   Anal. Biochem. 256:158-168(1998).
DR   EMBL; U34617; AAC36455.1; -.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1068 MW;  9A7598276728705A CRC64;


  Query Match             18.4%;  Score 25;  DB 13;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 DFGFPKHL 17
           | | || :
Db       3 DIGLPKRV 10


RESULT  12
O94554
ID   O94554      PRELIMINARY;      PRT;    16 AA.
AC   O94554;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CALMODULIN KINASE 2 (FRAGMENT).
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alemany V., Aligue R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57982; AAD09466.1; -.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1846 MW;  4A673B1F4C3288D9 CRC64;


  Query Match             18.4%;  Score 25;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 FGFP 14
           ||||
Db       5 FGFP 8


RESULT  13
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;


  Query Match             18.4%;  Score 25;  DB 13;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 2.1e+03;
  Matches    6;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       6 LLQMDFGFPKHLLVD 20
           ||    | |  | :|
Db       1 LLAQGLGXPTALALD 15


RESULT  14
Q9TWH5
ID   Q9TWH5      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   40 KDA GAP JUNCTION PROTEIN (FRAGMENT).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Heliothis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95347000.
RA   Ryerse J.S.;
RT   "Immunocytochemical, electrophoresis, and immunoblot analysis of
RT   Heliothis virescens gap junctions isolated in the presence and absence
RT   of protease inhibitors.";
RL   Cell Tissue Res. 281:179-186(1995).
SQ   SEQUENCE   20 AA;  2304 MW;  A298D3EB3E89586B CRC64;


  Query Match             18.4%;  Score 25;  DB 5;  Length 20;
  Best Local Similarity   29.4%;  Pred. No. 2.2e+03;
  Matches    5;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 LLLQMDFGFPKHLLVDF 21
           ::  :| |: : |  ||
Db       3 VIFNIDGGYLEFLTRDF 19


RESULT  15
Q9TTG3
ID   Q9TTG3      PRELIMINARY;      PRT;    20 AA.
AC   Q9TTG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GALACTOCEREBROSIDASE (FRAGMENT).
GN   CALC.
OS   Ateles belzebuth chamek.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ateles.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Canavez F.C.;
RT   "Gene mapping in Ateles paniscus chamek.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF099175; AAF21849.1; -.
DR   INTERPRO; IPR001286; -.
DR   PFAM; PF02057; Glyco_hydro_59; 1.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2317 MW;  EA3066655966B58F CRC64;


  Query Match             18.4%;  Score 25;  DB 6;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 2.2e+03;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;
```

```
  Query Match             20.6%;  Score 28;  DB 2;  Length 21;
  Best Local Similarity   45.5%;  Pred. No. 8e+02;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGFPK 15
          :| ::||  ||
Db      3 ILKKLDFSDPK 13


RESULT   7
Q9UGS1
ID   Q9UGS1      PRELIMINARY;      PRT;    12 AA.
AC   Q9UGS1;
DT   01-MAY-2000 (TrEMBLrel: 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ796I17.4 (NOVEL PROTEIN SIMILAR TO GS2) (FRAGMENT).
GN   DJ796I17.4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035398; CAB63074.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   12 AA;  1430 MW;  AF7740ABECB69AA6 CRC64;


  Query Match             19.9%;  Score 27;  DB 4;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 6.4e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGF 13
          ::|:||  |
Db      1 MILEMDMSF 9


RESULT   8
Q9TRR6
ID   Q9TRR6      PRELIMINARY;      PRT;    19 AA.
AC   Q9TRR6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN, CAP50=CA2+/PHOSPHOLIPID-BINDING PROTEIN
DE   L-14 FRAGMENT.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92250478.
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
SQ   SEQUENCE   19 AA;  2018 MW;  9A54062504B8322E CRC64;


  Query Match             19.9%;  Score 27;  DB 6;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 1e+03;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     11 FGFPKHLLVDFLQSLS 26
          ||  :  ::| | | |
Db      2 FGTDEQAIIDXLGSRS 17


RESULT   9
P70861
ID   P70861      PRELIMINARY;      PRT;    21 AA.
AC   P70861;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   THDF (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE; 97312006.
RA   Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT   "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT   sigma 70-like promoter.";
RL   Microbiology 143:1681-1690(1997).
DR   EMBL; U62901; AAB62742.1; -.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2432 MW;  F33E1EC548BD5B33 CRC64;


  Query Match             19.1%;  Score 26;  DB 2;  Length 21;
  Best Local Similarity   40.0%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      9 MDFGFPKHLL 18
          ::| ||: :|
Db      7 IEFDFPEGIL 16


RESULT  10
P82401
ID   P82401      PRELIMINARY;      PRT;    25 AA.
AC   P82401;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AUREIN 5.1.
OS   Litoria raniformis, and Litoria aurea (Australian frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
RN   [1]
RP   SEQUENCE.
RA   Rozek T., Wegener K.L., Bowie J.H., Olver I.N., Carver J.A.,
RA   Wallace J.C., Tyler M.J.;
RT   "The antibiotic and anticancer aurein peptides from the Australian
RT   bell frogs Litoria aurea and Litoria raniformis.";
RL   Eur. J. Biochem. 0:0-0(2000).
CC   -!- FUNCTION: HAS NO ANTIMICROBIAL OR ANTICANCER ACTIVITY.
SQ   SEQUENCE   25 AA;  2547 MW;  15C6169CD98AFC27 CRC64;


  Query Match             19.1%;  Score 26;  DB 13;  Length 25;
  Best Local Similarity   38.9%;  Pred. No. 2e+03;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPKHLLVDFLQ 23
          || :  |   :|:|| |:
Db      2 LLDIVTGLLGNLIVDVLK 19


RESULT  11
O73594
ID   O73594      PRELIMINARY;      PRT;    10 AA.
AC   O73594;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ZAX-2 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
```

```
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96092035.
RA   Linder D., Freund R., Kadenbach B.;
RT   "Species-specific expression of cytochrome c oxidase isozymes.";
RL   Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 112:461-469(1995).
SQ   SEQUENCE   24 AA;  2795 MW;  D49D27C03B61F803 CRC64;


  Query Match             24.3%;  Score 33;  DB 6;  Length 24;
  Best Local Similarity   58.3%;  Pred. No. 1.5e+02;
  Matches    7;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       6 LLQMDFGFPKHL 17
           | | | | | ||
Db      11 LFQEDNGIPVHL 22


RESULT   3
O31365
ID   O31365      PRELIMINARY;      PRT;    13 AA.
AC   O31365;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP90;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93701; AAC45536.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1382 MW;  08EFB4180E0BC720 CRC64;


  Query Match             23.5%;  Score 32;  DB 2;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CLLLQMDFG 12
           | :| :|||
Db       4 CAILVLDFG 12


RESULT   4
Q9ZG55
ID   Q9ZG55      PRELIMINARY;      PRT;    21 AA.
AC   Q9ZG55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ATP-BINDING PROTEIN (FRAGMENT).
GN   RECF.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087306; AAD04082.1; -.
KW   ATP-binding.
FT   NON_TER       1      1
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2336 MW;  0185D9AC428276D9 CRC64;


  Query Match             22.4%;  Score 30.5;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 3.3e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       6 LLQMDFGFPKHL 17
           :||:   ||||:
Db       8 ILQL-ISFPKHI 18


RESULT   5
Q46081
ID   Q46081      PRELIMINARY;      PRT;    24 AA.
AC   Q46081;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NITROGENASE 3 SUBUNIT H (FRAGMENT).
GN   ANFH.
OS   Clostridium hungatei.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC   Clostridium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B3B;
RA   Chen T., Leschine S.B.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U59415; AAB02934.1; -.
DR   INTERPRO; IPR000392; -.
DR   PFAM; PF00142; fer4_NifH; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   24 AA;  2752 MW;  C014801651E11BFC CRC64;


  Query Match             22.1%;  Score 30;  DB 2;  Length 24;
  Best Local Similarity   37.5%;  Pred. No. 4.5e+02;
  Matches    6;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      10 DFGFPKHLLVDFLQSL 25
           :|  || | :| |: :
Db       1 EFVIPKPLTMDQLEDM 16


RESULT   6
Q9X3D0
ID   Q9X3D0      PRELIMINARY;      PRT;    21 AA.
AC   Q9X3D0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
GN   PETD.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070133; AAD20742.1; -.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2410 MW;  CC947A212386FE23 CRC64;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:40:02 ; Search time 28.23 Seconds
(without alignments)
85.998 Million cell updates/sec

Title: US-08-934-367-50
Perfect score: 136
Sequence: 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 6467

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	24.3	13	2	O31296	O31296 borrelia af
2	33	24.3	24	6	Q9TR30	Q9tr30 ovis aries
3	32	23.5	13	2	O31365	O31365 borrelia ga
4	30.5	22.4	21	2	Q9ZG55	Q9zg55 chlamydia t
5	30	22.1	24	2	Q46081	Q46081 clostridium
6	28	20.6	21	2	Q9X3D0	Q9x3d0 prochloroco
7	27	19.9	12	4	Q9UGS1	Q9ugs1 homo sapien
8	27	19.9	19	6	Q9TRR6	Q9trr6 oryctolagus
9	26	19.1	21	2	P70861	P70861 borrelia bu
10	26	19.1	25	13	P82401	P82401 litoria ran
11	25	18.4	10	13	O73594	O73594 gallus gall
12	25	18.4	16	3	O94554	O94554 schizosacch
13	25	18.4	19	13	Q9PS70	Q9ps70 gallus gall
14	25	18.4	20	5	Q9TWH5	Q9twh5 heliothis v
15	25	18.4	20	6	Q9TTG3	Q9ttg3 ateles belz
16	25	18.4	21	4	Q16575	Q16575 homo sapien
17	25	18.4	23	6	Q9TRZ8	Q9trz8 canis famil
18	25	18.4	24	12	Q69137	Q69137 human herpe
19	25	18.4	25	4	Q9UQU9	Q9uqu9 homo sapien
20	25	18.4	26	2	Q9R6Z1	Q9r6z1 photobacter
21	24.5	18.0	17	4	Q9UJH1	Q9ujh1 homo sapien
22	24.5	18.0	22	13	Q9PS42	Q9ps42 gallus gall
23	24.5	18.0	25	13	Q9PS41	Q9ps41 gallus gall
24	24	17.6	12	11	Q64296	Q64296 mus musculu
25	24	17.6	13	2	O31364	O31364 borrelia ga
26	24	17.6	13	2	O34622	O34622 borrelia bu
27	24	17.6	15	11	Q9Z1I4	Q9z1i4 mus musculu
28	24	17.6	19	5	Q9W5Q8	Q9w5q8 drosophila
29	24	17.6	22	12	Q85460	Q85460 avian myelo
30	24	17.6	25	2	O67969	O67969 prochloroco
31	23.5	17.3	15	2	Q9R4U7	Q9r4u7 acinetobact
32	23	16.9	8	2	O85406	O85406 coxiella bu
33	23	16.9	9	5	O96417	O96417 drosophila
34	23	16.9	10	2	Q9X534	Q9x534 leclercia a
35	23	16.9	10	2	Q9X533	Q9x533 escherichia
36	23	16.9	13	2	O34770	O34770 borrelia af
37	23	16.9	13	11	O35758	O35758 rattus norv
38	23	16.9	14	6	Q9TRQ7	Q9trq7 bos taurus
39	23	16.9	17	2	Q9X3I3	Q9x3i3 prochloroco
40	23	16.9	18	13	P82068	P82068 litoria gen
41	23	16.9	20	2	Q9R4F1	Q9r4f1 desulfovibr
42	23	16.9	21	4	Q9UMU6	Q9umu6 homo sapien
43	23	16.9	21	8	Q9ZYB7	Q9zyb7 spinaria sp
44	23	16.9	21	10	Q41496	Q41496 solanum tub
45	23	16.9	21	12	Q80817	Q80817 human t-cel

ALIGNMENTS

RESULT 1
O31296
ID O31296 PRELIMINARY; PRT; 13 AA.
AC O31296;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE GUAA (FRAGMENT).
GN GUAA.
OS Borrelia afzelii.
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP21;
RX MEDLINE; 97426044.
RA Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA Rosa P.;
RT "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT plasmid structure and targeted inactivation of the ospC gene.";
RL Mol. Microbiol. 25:361-374(1997).
DR EMBL; U93696; AAC45526.1; -.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1412 MW; 08EFA2C20E195720 CRC64;

Query Match 24.3%; Score 33; DB 2; Length 13;
Best Local Similarity 41.7%; Pred. No. 80;
Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LDGCLLLQMDFG 12
:: | :| :|||
Db 1 MNDCAILILDFG 12

RESULT 2
Q9TR30
ID Q9TR30 PRELIMINARY; PRT; 24 AA.
AC Q9TR30;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT VIIA-L (FRAGMENT).

Query Match 16.2%; Score 22; DB 1; Length 26;
Best Local Similarity 71.4%; Pred. No. 2.5e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      17 LLVDFLQ 23
           |||| |:
Db      13 LLVDVLK 19
```

Search completed: December 21, 2000, 08:40:18
Job time: 116 sec

```
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
CC       PHOTOSYSTEM II (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
CC       WITH THE PHOTOSYSTEM II COMPLEX (BY SIMILARITY).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC       (SPOT N179) IS: 5.9, ITS MW IS: 22 KDA.
CC   -!- SIMILARITY: TO OTHER OEE2 SUBUNITS.
KW   Photosynthesis; Photosystem II; Chloroplast; Thylakoid membrane.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1294 MW;  C6772B0D54D7C44D CRC64;


  Query Match             16.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 FGFPK 15
           || ||
Db       9 FGAPK 13


RESULT  13
CHH3_BOMMO
ID   CHH3_BOMMO     STANDARD;      PRT;     17 AA.
AC   P20729;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   CHORION CLASS HIGH-CYSTEINE HCA PROTEIN 13 PRECURSOR (HC-A.13)
DE   (FRAGMENT).
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85083111.
RA   Rodakis G.C., Lecanidou R., Eickbush T.H.;
RT   "Diversity in a chorion multigene family created by tandem
RT   duplications and a putative gene-conversion event.";
RL   J. Mol. Evol. 20:265-273(1984).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC       SILK MOTH.
CC   -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
CC       BELONG CLASSES A, CA AND HCA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01068; CAB57790.1; -.
DR   PIR; B23219; B23219.
KW   Eggshell; Chorion; Repeat; Multigene family; Signal.
FT   SIGNAL       1     >17
FT   NON_TER     17     17
SQ   SEQUENCE   17 AA;  1913 MW;  5E634508C5355C9C CRC64;


  Query Match             16.2%;  Score 22;  DB 1;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LDGCLL 6
            : |||:
Db      11 VQGCLI 16


RESULT  14
COXN_THUOB
ID   COXN_THUOB     STANDARD;      PRT;     20 AA.
AC   P80980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2303 MW;  0A33BD34006E5AA6 CRC64;


  Query Match             16.2%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 DFGFPKH 16
           | | | |
Db      14 DNGMPVH 20


RESULT  15
CT21_LITCI
ID   CT21_LITCI     STANDARD;      PRT;     26 AA.
AC   P81847; P81848; P81849; P81850;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CITROPIN 2.1.3 [CONTAINS: CITROPIN 2.1.2; CITROPIN 2.1.1; CITROPIN
DE   2.1].
OS   Litoria citropa (Australian blue mountains tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 99435977.
RA   Wegener K.L., Wabnitz P.A., Carver J.A., Bowie J.H., Chia B.C.S.,
RA   Wallace J.C., Tyler M.J.;
RT   "Host defence peptides from the skin glands of the Australian blue
RT   mountains tree-frog Litoria citropa. Solution structure of the
RT   antibacterial peptide citropin 1.1.";
RL   Eur. J. Biochem. 265:627-637(1999).
CC   -!- TISSUE SPECIFICITY: DORSAL AND SUBMENTAL SKIN GLANDS.
FT   PEPTIDE       1     22       CITROPIN 2.1.
FT   PEPTIDE       1     23       CITROPIN 2.1.1.
FT   PEPTIDE       1     25       CITROPIN 2.1.2.
FT   PEPTIDE       1     26       CITROPIN 2.1.3.
SQ   SEQUENCE   26 AA;  2519 MW;  7490C28EB2E4899A CRC64;
```

```
Qy      14 PKHLLV 19
             | ||||
Db       7 PFHLLV 12


RESULT   9
DCMM_PSECF
ID   DCMM_PSECF     STANDARD;      PRT;    14 AA.
AC   P19914;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0142; PL0142.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1756 MW;  65583C6D1FB7C25B CRC64;


  Query Match             17.6%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   23.1%;  Pred. No. 6.3e+02;
  Matches    3;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       5 LLLQMDFGFPKHL 17
           :: : ::  |||:
Db       2 MIPRFEYHAPKHV 14


RESULT  10
YCXD_ODOSI
ID   YCXD_ODOSI     STANDARD;      PRT;    26 AA.
AC   P49839;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HYPOTHETICAL 3.2 KDA PROTEIN IN RPOC2-RPS2 INTERGENIC REGION (ORF26B).
OS   Odontella sinensis.
OG   Chloroplast.
OC   Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC   Biddulphiophycidae; Eupodiscales; Eupodiscaceae; Odontella.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT   "The chloroplast genome of a chlorophyll a+c-containing alga,
RT   Odontella sinensis.";
RL   Plant Mol. Biol. Rep. 13:336-342(1995).
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; Z67753; CAA91747.1; -.
KW   Chloroplast; Hypothetical protein.
SQ   SEQUENCE   26 AA;  3137 MW;  8305B587C0A30B36 CRC64;


  Query Match             17.3%;  Score 23.5;  DB 1;  Length 26;
  Best Local Similarity   60.0%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    1;  Mismatches    0;  Indels    3;  Gaps    1;

Qy       1 LDG---CLLL 7
             |||   |||:
Db      10 LDGFIECLLM 19


RESULT  11
DHE2_THUTH
ID   DHE2_THUTH     STANDARD;      PRT;    16 AA.
AC   P20016;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   NAD-SPECIFIC GLUTAMATE DEHYDROGENASE (EC 1.4.1.2) (NAD-GDH)
DE   (FRAGMENT).
OS   Thunnus thynnus (Bluefin tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LIVER;
RX   MEDLINE; 76253947.
RA   Veronese F.M., Bevilacqua R., Boccu E., Brown D.M.;
RT   "Purification, characteristics and sequence of a peptide containing
RT   an essential lysine residue.";
RL   Biochim. Biophys. Acta 445:1-13(1976).
CC   -!- CATALYTIC ACTIVITY: L-GLUTAMATE + H(2)O + NAD(+) = 2-OXOGLUTARATE
CC       + NH(3) + NADH.
CC   -!- SIMILARITY: BELONGS TO THE GLU/LEU/PHE/VAL DEHYDROGENASES FAMILY.
DR   PIR; A12729; A12729.
DR   INTERPRO; IPR001625; -.
DR   PROSITE; PS00074; GLFV_DEHYDROGENASE; PARTIAL.
KW   Oxidoreductase; NAD.
FT   NON_TER       1      1
FT   ACT_SITE     12     12
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1518 MW;  FF299AA7C5F1062F CRC64;


  Query Match             16.9%;  Score 23;  DB 1;  Length 16;
  Best Local Similarity   33.3%;  Pred. No. 1e+03;
  Matches    4;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       4 CLLLQMDFGFPK 15
           | :: : ||  |
Db       1 CAVVDVPFGGAK 12


RESULT  12
PSBP_PINPS
ID   PSBP_PINPS     STANDARD;      PRT;    13 AA.
AC   P81668;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   OXYGEN-EVOLVING ENHANCER PROTEIN 2 (OEE2) (23 KDA SUBUNIT OF OXYGEN
DE   EVOLVING SYSTEM OF PHOTOSYSTEM II) (FRAGMENT).
GN   PSBP.
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferopsida; Coniferales; Pinaceae; Pinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=NEEDLE;
RX   MEDLINE; 99274088.
```

```
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
DR   PIR; S43632; S43632.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      23     23
SQ   SEQUENCE   23 AA;  2635 MW;  BCBED43FBAD9C509 CRC64;


  Query Match             19.9%;  Score 27;  DB 1;  Length 23;
  Best Local Similarity   50.0%;  Pred. No. 3.6e+02;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPKHL 17
          | |   | | ||
Db     11 LFQAXNGIPVHL 22


RESULT   6
COXK_CANFA
ID   COXK_CANFA     STANDARD;      PRT;    23 AA.
AC   Q9TRZ8;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-HEART, MITOCHONDRIAL
DE   (EC 1.9.3.1) (CYTOCHROME C OXIDASE SUBUNIT VIIA-H) (COX VIIA-M)
DE   (FRAGMENT).
GN   COX7A1 OR COX7AH.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BRAIN, AND HEART;
RX   MEDLINE; 96092035.
RA   Linder D., Freund R., Kadenbach B.;
RT   "Species-specific expression of cytochrome c oxidase isozymes.";
RL   Comp. Biochem. Physiol. 112B:461-469(1995).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
KW   Oxidoreductase; Inner membrane; Mitochondrion.
FT   NON_TER      23     23
SQ   SEQUENCE   23 AA;  2627 MW;  AD3EA34B61FF73CE CRC64;


  Query Match             18.4%;  Score 25;  DB 1;  Length 23;
  Best Local Similarity   50.0%;  Pred. No. 7.4e+02;
  Matches    6;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      6 LLQMDFGFPKHL 17
          | | | | |  |
Db     11 LFQADNGLPVXL 22


RESULT   7
CH60_DROME
ID   CH60_DROME     STANDARD;      PRT;    11 AA.
AC   P35380;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MITOCHONDRIAL MATRIX PROTEIN P1 (60 KDA CHAPERONIN) (HEAT SHOCK
DE   PROTEIN 60) (HSP-60) (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).
GN   MMP-P1 OR HSP60.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VALLECAS; TISSUE=WING IMAGINAL DISK;
RX   MEDLINE; 93272852.
RA   Santaren J.F., van Damme J., Puype M., Vandekerckhove J.,
RA   Garcia-Bellido A.;
RT   "Identification of Drosophila wing imaginal disc proteins by two-
RT   dimensional gel analysis and microsequencing.";
RL   Exp. Cell Res. 206:220-226(1993).
CC   -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
CC       MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC       IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC       REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC       UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR   FLYBASE; FBgn0010375; Mmp-P1.
DR   INTERPRO; IPR001844; -.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW   Chaperone; ATP-binding; Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1243 MW;  78501A366365A6DB CRC64;


  Query Match             17.6%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   27.3%;  Pred. No. 4.9e+02;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGFPK 15
          ::::  :| ||
Db      1 VIIEQSWGSPK 11


RESULT   8
ANGT_HORSE
ID   ANGT_HORSE     STANDARD;      PRT;    14 AA.
AC   P01016;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   ANGIOTENSINOGEN (FRAGMENT).
GN   AGT.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RA   Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
RT   "The preparation, purification, and amino acid sequence of a
RT   polypeptide renin substrate.";
RL   J. Exp. Med. 106:439-453(1957).
CC   -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC       CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC       CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC       PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC       PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC       BALANCE OF BODY FLUIDS.
CC   -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   PIR; A01250; A01250.
DR   INTERPRO; IPR000215; -.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Vasoconstrictor; Plasma; Serpin.
FT   PEPTIDE       1     10       ANGIOTENSIN I.
FT   PEPTIDE       1      8       ANGIOTENSIN II.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1759 MW;  2E9921F8EEEFBDD7 CRC64;


  Query Match             17.6%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 6.3e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
RX   MEDLINE; 76081726.
RA   Birken S., Wilner G.D., Canfield R.E.;
RT   "Studies of the structure of canine fibrinogen.";
RL   Thromb. Res. 7:599-610(1975).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC       CONVERTED INTO THE HARD CLOT BY FACTOR XIIIA WHICH CATALYZES THE
CC       EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC       (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC       MONOMERS.
DR   PIR; A05298; A05298.
DR   INTERPRO; IPR002181; -.
DR   PROSITE; PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
KW   Blood coagulation; Plasma.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2690 MW;  D427050783F52628 CRC64;


  Query Match             20.6%;  Score 28;  DB 1;  Length 24;
  Best Local Similarity   45.5%;  Pred. No. 2.6e+02;
  Matches    5;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 DGCLLLQMDFG 12
           | | :|   ||.
Db       6 DNCCILDERFG 16


RESULT   3
IRBP_SHEEP
ID   IRBP_SHEEP     STANDARD;      PRT;     24 AA.
AC   P12663;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   INTERPHOTORECEPTOR RETINOID-BINDING PROTEIN (IRBP) (INTERSTITIAL
DE   RETINOL-BINDING PROTEIN) (FRAGMENT).
GN   RBP3.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 86301171.
RA   Fong S.L., Cook R.G., Alvarez R.A., Liou G.I., Landers R.A.,
RA   Bridges C.D.B.;
RT   "N-terminal sequence homologies in interstitial retinol-binding
RT   proteins from 10 vertebrate species.";
RL   FEBS Lett. 205:309-312(1986).
CC   -!- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOIDS BETWEEN
CC       THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL
CC       PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.
CC   -!- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES
CC       THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT
CC       EPITHELIUM CELLS.
DR   PIR; A24417; A24417.
KW   Vitamin A; Transport.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2799 MW;  02EDEBE61A8E4523 CRC64;


  Query Match             20.6%;  Score 28;  DB 1;  Length 24;
  Best Local Similarity   33.3%;  Pred. No. 2.6e+02;
  Matches    6;  Conservative    6;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 LDGCLLLQMDFGFPKHLL 18
           ||   :|  :: ||::|:
Db       7 LDMAQVLLDNYTFPENLM 24


RESULT   4
IRBP_PIG
ID   IRBP_PIG       STANDARD;      PRT;     25 AA.
AC   P12662;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   INTERPHOTORECEPTOR RETINOID-BINDING PROTEIN (IRBP) (INTERSTITIAL
DE   RETINOL-BINDING PROTEIN) (FRAGMENT).
GN   RBP3.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 86301171.
RA   Fong S.L., Cook R.G., Alvarez R.A., Liou G.I., Landers R.A.,
RA   Bridges C.D.B.;
RT   "N-terminal sequence homologies in interstitial retinol-binding
RT   proteins from 10 vertebrate species.";
RL   FEBS Lett. 205:309-312(1986).
CC   -!- FUNCTION: IRBP SHUTTLES 11-CIS AND ALL TRANS RETINOIDS BETWEEN
CC       THE RETINOL ISOMERASE IN THE PIGMENT EPITHELIUM AND THE VISUAL
CC       PIGMENTS IN THE PHOTORECEPTOR CELLS OF THE RETINA.
CC   -!- SUBCELLULAR LOCATION: INTERPHOTORECEPTOR MATRIX THAT PERMEATES
CC       THE SPACE BETWEEN THE RETINA AND THE CONTIGUOUS LAYER OF PIGMENT
CC       EPITHELIUM CELLS.
DR   PIR; B24417; B24417.
KW   Vitamin A; Transport.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2813 MW;  4E751DFA160231B7 CRC64;


  Query Match             20.6%;  Score 28;  DB 1;  Length 25;
  Best Local Similarity   33.3%;  Pred. No. 2.7e+02;
  Matches    6;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LDGCLLLQMDFGFPKHLL 18
           ||   :|  :: ||: |:
Db       7 LDTAKILLDNYTFPESLM 24


RESULT   5
COXJ_ONCMY
ID   COXJ_ONCMY     STANDARD;      PRT;     23 AA.
AC   P80333;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-LIVER (EC 1.9.3.1) (VIIIC)
DE   (FRAGMENT).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LIVER;
RX   MEDLINE; 94237150.
RA   Freund R., Kadenbach B.;
RT   "Identification of tissue-specific isoforms for subunits Vb and VIIa
RT   of cytochrome c oxidase isolated from rainbow trout.";
RL   Eur. J. Biochem. 221:1111-1116(1994).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:40:18 ; Search time 12.36 Seconds
(without alignments)
67.204 Million cell updates/sec

Title: US-08-934-367-50
Perfect score: 136
Sequence: 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1478

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	24.3	24	1	COXJ_SHEEP	Q9tr30 ovis aries
2	28	20.6	24	1	FIBG_CANFA	P12800 canis famil
3	28	20.6	24	1	IRBP_SHEEP	P12663 ovis aries
4	28	20.6	25	1	IRBP_PIG	P12662 sus scrofa
5	27	19.9	23	1	COXJ_ONCMY	P80333 oncorhynchu
6	25	18.4	23	1	COXK_CANFA	Q9trz8 canis famil
7	24	17.6	11	1	CH60_DROME	P35380 drosophila
8	24	17.6	14	1	ANGT_HORSE	P01016 equus cabal
9	24	17.6	14	1	DCMM_PSECF	P19914 pseudomonas
10	23.5	17.3	26	1	YCXD_ODOSI	P49839 odontella s
11	23	16.9	16	1	DHE2_THUTH	P20016 thunnus thy
12	22	16.2	13	1	PSBP_PINPS	P81668 pinus pinas
13	22	16.2	17	1	CHH3_BOMMO	P20729 bombyx mori
14	22	16.2	20	1	COXN_THUOB	P80980 thunnus obe
15	22	16.2	26	1	CT21_LITCI	P81847 litoria cit
16	21.5	15.8	25	1	ALR_PSEFL	P33162 pseudomonas
17	21	15.4	9	1	SAMP_MUSCA	P19095 mustelus ca
18	21	15.4	23	1	IRBP_RABIT	P12664 oryctolagus
19	21	15.4	24	1	CR17_LITXA	P56232 litoria xan
20	21	15.4	24	1	CR18_LITCH	P81251 litoria chl
21	21	15.4	25	1	AMP3_MELGA	P80393 meleagris g
22	21	15.4	26	1	AMD1_CHICK	P81073 gallus gall
23	21	15.4	26	1	AMD1_RABIT	P81072 oryctolagus
24	21	15.4	26	1	MEL_APIDO	P01502 apis dorsat
25	20	14.7	11	1	TEML_RANTE	P56923 rana tempor
26	20	14.7	16	1	MDH_SYNY4	P80460 synechocyst
27	20	14.7	18	1	DRPH_UCAPU	P08871 uca pugilat
28	20	14.7	21	1	ATPB_PHYPA	P80658 physcomitre
29	20	14.7	21	1	OMP1_ACTPL	P80368 actinobacil
30	20	14.7	21	1	TRYP_PROAT	P35051 protopterus
31	20	14.7	23	1	SODM_RANCA	P36215 rana catesb
32	20	14.7	24	1	CCAA_STRTI	P80436 streptomyce
33	19	14.0	10	1	GSI5_BACSU	P80861 bacillus su
34	19	14.0	12	1	CD11_LITXA	P56245 litoria xan
35	19	14.0	12	1	CD14_LITXA	P56246 litoria xan
36	19	14.0	13	1	SODM_CANFA	P54712 canis famil
37	19	14.0	19	1	MDH_SHEPU	P82177 shewanella
38	19	14.0	19	1	UP21_UPEIN	P82027 uperoleia i
39	19	14.0	19	1	UP25_UPEIN	P82031 uperoleia i
40	19	14.0	20	1	ITRA_ALBJU	P24925 albizzia ju
41	19	14.0	21	1	DCMS_PSECA	P19921 pseudomonas
42	19	14.0	24	1	FEDG_AMYME	P80707 amycolatops
43	19	14.0	24	1	LEC_CROST	P16351 crotalaria
44	19	14.0	24	1	PEP1_ASPNG	P55749 aspergillus
45	19	14.0	25	1	BOTR_BOTJA	P22028 bothrops ja

ALIGNMENTS

RESULT 1
COXJ_SHEEP
ID COXJ_SHEEP STANDARD; PRT; 24 AA.
AC Q9TR30;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-LIVER/HEART, MITOCHONDRIAL
DE (EC 1.9.3.1) (CYTOCHROME C OXIDASE SUBUNIT VIIA-L) (FRAGMENT).
GN COX7A2 OR COX7AL.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Ovis.
RN [1]
RP SEQUENCE.
RC TISSUE=LIVER, AND HEART;
RX MEDLINE; 96092035.
RA Linder D., Freund R., Kadenbach B.;
RT "Species-specific expression of cytochrome c oxidase isozymes.";
RL Comp. Biochem. Physiol. 112B:461-469(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 24 24
SQ SEQUENCE 24 AA; 2795 MW; D49D27C03B61F803 CRC64;

Query Match 24.3%; Score 33; DB 1; Length 24;
Best Local Similarity 58.3%; Pred. No. 43;
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 6 LLQMDFGFPKHL 17
| | | | | ||
Db 11 LFQEDNGIPVHL 22

RESULT 2
FIBG_CANFA
ID FIBG_CANFA STANDARD; PRT; 24 AA.
AC P12800;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE FIBRINOGEN GAMMA-CHAIN (FRAGMENT).
GN FGG.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN [1]
RP SEQUENCE.

```
RESULT  14
PC4030
rRNA endonuclease (EC 3.1.27.10) - shiitake mushroom (fragment)
N;Alternate names: nuclease Le3
C;Species: Lentinula edodes (shiitake mushroom)
C;Date: 21-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 20-Mar-1996
C;Accession: PC4030
R;Kobayashi, H.; Inokuchi, N.; Koyama, T.; Tomita, M.; Irie, M.
Biosci. Biotechnol. Biochem. 59, 1169-1171, 1995
A;Title: Purification and characterization of the 2nd 5'-nucleotide-forming nuclease fro
A;Reference number: PC4030; MUID:95337563
A;Accession: PC4030
A;Molecule type: DNA
A;Residues: 1-23 <KOB>
C;Comment: This enzyme has 3'-nucleotidase activity.
C;Keywords: endonuclease; hydrolase


  Query Match             18.4%;  Score 25;  DB 2;  Length 23;
  Best Local Similarity   42.9%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      11 FGFPKHLLVDFLQS 24
           :|   | || |: |
Db       1 WGMLGHELVGFIAS 14


RESULT  15
B61597
cytochrome P450 AL-2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: B61597
R;Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.
Drug Metab. Dispos. 19, 291-297, 1991
A;Title: Purification and aminopyrine monooxygenase activity of liver microsomal cytochr
A;Reference number: A61597; MUID:91292910
A;Accession: B61597
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <SHI>


  Query Match             18.0%;  Score 24.5;  DB 2;  Length 14;
  Best Local Similarity   47.1%;  Pred. No. 9.7e+02;
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    5;  Gaps    1;

Qy       5 LLLQMDFGFPKHLLVDF 21
           ||| : |     ||| |
Db       3 LLLSLSF-----LLVGF 14


Search completed: December 21, 2000, 08:39:30
Job time: 134 sec
```

C;Superfamily: mammalian cytochrome-c oxidase chain VIIa
C;Keywords: liver; membrane-associated complex; mitochondrion; oxidoreductase

Query Match 19.9%; Score 27; DB 2; Length 23;
Best Local Similarity 50.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy       6 LLQMDFGFPKHL 17
           | |   | | ||
Db      11 LFQAXNGIPVHL 22
```

RESULT 9
PC2199
alicyclic amine N-sulfotransferase (EC 2.8.2.-) - rat (fragment)
N;Alternate names: NST-1
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Sep-1994 #sequence_revision 05-Apr-1995 #text_change 03-May-1996
C;Accession: PC2199
R;Naritomi, Y.; Niwa, T.; Shiraga, T.; Iwasaki, K.; Noda, K.
Biol. Pharm. Bull. 17, 1008-1011, 1994
A;Title: Isolation and characterization of an alicyclic amine N-sulfotransferase from fe
A;Reference number: PC2199; MUID:95119961
A;Accession: PC2199
A;Molecule type: protein
A;Residues: 1-24 <NAR>
A;Experimental source: liver
C;Comment: This enzyme catalyzes sulfonation not only of the alicyclic amine but also de
C;Superfamily: alcohol sulfotransferase
C;Keywords: sulfotransferase

Query Match 19.9%; Score 27; DB 2; Length 24;
Best Local Similarity 62.5%; Pred. No. 7e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      11 FGFPKHLL 18
           || ||  |
Db      14 FGIPKETL 21
```

RESULT 10
S59492
formate dehydrogenase alpha chain - Alcaligenes eutrophus (fragment)
C;Species: Alcaligenes eutrophus
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59492
R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.
Biol. Chem. Hoppe-Seyler 376, 561-568, 1995
A;Title: Structural and immunological studies on the soluble formate dehydrogenase from
A;Reference number: S59492; MUID:96145736
A;Accession: S59492
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRI>

Query Match 19.1%; Score 26; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 6.1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy       8 QMDFGFP 14
           ::|||  |
Db       6 EIDFGTP 12
```

RESULT 11
PX0078
alanine dehydrogenase (EC 1.4.1.1) - oscillatoriacean cyanobacterium (fragment)
C;Species: oscillatoriacean cyanobacterium
C;Date: 22-Apr-1995 #sequence_revision 26-May-1995 #text_change 07-May-1999
C;Accession: PX0078
R;Sawa, Y.; Tani, M.; Murata, K.; Shibata, H.; Ochiai, H.
J. Biochem. 116, 995-1000, 1994
A;Title: Purification and characterization of alanine dehydrogenase from a cyanobacte
A;Reference number: PX0078; MUID:95204408
A;Accession: PX0078
A;Molecule type: protein
A;Residues: 1-21 <SAW>
A;Note: the source is designated as Phormidium lapideum
C;Comment: This enzyme catalyzes a reversible oxidative deamination of L-alanine to p
C;Superfamily: alanine dehydrogenase; alanine dehydrogenase homology
C;Keywords: oxidoreductase

Query Match 19.1%; Score 26; DB 2; Length 21;
Best Local Similarity 44.4%; Pred. No. 8.7e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       9 MDFGFPKHL 17
           |: | || :
Db       1 MEIGVPKEI 9
```

RESULT 12
T44936
calmodulin kinase 2 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44936
R;Alemany, V.; Algue, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z22873
A;Accession: T44936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-16 <ALE>
A;Cross-references: EMBL:U57982; PIDN:AAD09466.1

Query Match 18.4%; Score 25; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 9.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      11 FGFP 14
           ||||
Db       5 FGFP 8
```

RESULT 13
A60743
ornithine carbamoyltransferase (EC 2.1.3.3), anabolic - Aeromonas formicans (fragment
C;Species: Aeromonas formicans
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: A60743
R;Tricot, C.; De Coen, J.L.; Momin, P.; Falmagne, P.; Stalon, V.
J. Gen. Microbiol. 135, 2453-2464, 1989
A;Title: Evolutionary relationships among bacterial carbamoyltransferases.
A;Reference number: A60743; MUID:90188287
A;Accession: A60743
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-17 <TRI>
C;Keywords: transferase

Query Match 18.4%; Score 25; DB 2; Length 17;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      15 KHLLVD 20
           |||| |
Db       1 KHLLKD 6
```

A05298
fibrinogen gamma chain - dog (fragment)
C;Species: Canis lupus familiaris (dog)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 10-Mar-1994
C;Accession: A05298
R;Birken, S.; Wilner, G.D.; Canfield, R.E.
Thromb. Res. 7, 599-610, 1975
A;Title: Studies of the structure of canine fibrinogen.
A;Reference number: A94308; MUID:76081726
A;Accession: A05298
A;Molecule type: protein
A;Residues: 1-24 <BIR>
C;Superfamily: fibrinogen gamma chain; fibrinogen beta/gamma homology
C;Keywords: blood coagulation; liver; plasma

Query Match 20.6%; Score 28; DB 2; Length 24;
Best Local Similarity 45.5%; Pred. No. 4.9e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 DGCLLLQMDFG 12
| | :| ||
Db 6 DNCCILDERFG 16

RESULT 4
A24417
interphotoreceptor retinoid-binding protein - sheep (fragment)
N;Alternate names: interstitial retinol-binding protein
C;Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C;Accession: A24417
R;Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A;Title: N-terminal sequence homologies in interstitial retinol-binding proteins from 1
A;Reference number: A91365; MUID:86301171
A;Accession: A24417
A;Molecule type: protein
A;Residues: 1-24 <FON>
C;Superfamily: interphotoreceptor retinoid-binding protein
C;Keywords: duplication

Query Match 20.6%; Score 28; DB 2; Length 24;
Best Local Similarity 33.3%; Pred. No. 4.9e+02;
Matches 6; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Qy 1 LDGCLLLQMDFGFPKHLL 18
|| :| :: ||::|:
Db 7 LDMAQVLLDNYTFPENLM 24

RESULT 5
B24417
interphotoreceptor retinoid-binding protein - pig (fragment)
N;Alternate names: interstitial retinol-binding protein
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 21-Nov-1997
C;Accession: B24417
R;Fong, S.L.; Cook, R.G.; Alvarez, R.A.; Liou, G.I.; Landers, R.A.; Bridges, C.D.B.
FEBS Lett. 205, 309-312, 1986
A;Title: N-terminal sequence homologies in interstitial retinol-binding proteins from 1
A;Reference number: A91365; MUID:86301171
A;Accession: B24417
A;Molecule type: protein
A;Residues: 1-25 <FON>
C;Superfamily: interphotoreceptor retinoid-binding protein
C;Keywords: duplication

Query Match 20.6%; Score 28; DB 2; Length 25;
Best Local Similarity 33.3%; Pred. No. 5.1e+02;
Matches 6; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LDGCLLLQMDFGFPKHLL 18
|| :| :: ||: |:
Db 7 LDTAKILLDNYTFPESLM 24

RESULT 6
S36662
dermorphin (Lys-7) [validated] - two-colored leaf frog
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 10-Dec-1993 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000
C;Accession: S36662
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil,
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of
A;Reference number: S21152; MUID:92339502
A;Accession: S36662
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>

Query Match 19.9%; Score 27; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FGFPK 15
||:||
Db 3 FGYPK 7

RESULT 7
A32521
hexokinase (EC 2.7.1.1) I peptide III - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C;Accession: A32521
R;Schirch, D.M.; Wilson, J.E.
Arch. Biochem. Biophys. 257, 1-12, 1987
A;Title: Rat brain hexokinase: amino acid sequence at the substrate hexose binding si
A;Reference number: A90080; MUID:87324917
A;Accession: A32521
A;Molecule type: protein
A;Residues: 1-21 <SCH>
C;Superfamily: human hexokinase I; hexokinase homology
C;Keywords: ATP; glycolysis; phosphotransferase

Query Match 19.9%; Score 27; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 7 LQMDFGFPKH 16
| | || |
Db 3 LGFTFSFPXH 12

RESULT 8
S43632
cytochrome-c oxidase (EC 1.9.3.1) chain VIIa, hepatic - rainbow trout (fragment)
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 20-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: S43632
R;Freund, R.; Kadenbach, B.
Eur. J. Biochem. 221, 1111-1116, 1994
A;Title: Identification of tissue-specific isoforms for subunits Vb and VIIa of cytoc
A;Reference number: S43624; MUID:94237150
A;Accession: S43632
A;Molecule type: protein
A;Residues: 1-23 <FRE>
A;Note: the source is designated as Salmo gairdneri
C;Genetics:
A;Genome: nuclear

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:39:30 ; Search time 19.49 Seconds
(without alignments)
84.657 Million cell updates/sec

Title: US-08-934-367-50
Perfect score: 136
Sequence: 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 5255

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	21.3	26	2	B59018	MUC1 enhancer bind
2	28	20.6	19	2	A39504	octamer-binding pr
3	28	20.6	24	2	A05298	fibrinogen gamma c
4	28	20.6	24	2	A24417	interphotoreceptor
5	28	20.6	25	2	B24417	interphotoreceptor
6	27	19.9	7	2	S36662	dermorphin (Lys-7)
7	27	19.9	21	2	A32521	hexokinase (EC 2.7
8	27	19.9	23	2	S43632	cytochrome-c oxida
9	27	19.9	24	2	PC2199	alicyclic amine N-
10	26	19.1	15	2	S59492	formate dehydrogen
11	26	19.1	21	2	PX0078	alanine dehydrogen
12	25	18.4	16	2	T44936	calmodulin kinase
13	25	18.4	17	2	A60743	ornithine carbamoy
14	25	18.4	23	2	PC4030	rRNA endonuclease
15	24.5	18.0	14	2	B61597	cytochrome P450 AL
16	24	17.6	13	2	S03879	6-phosphofructokin
17	24	17.6	14	2	A01250	angiotensin precur
18	24	17.6	14	2	PL0142	carbon-monoxide de
19	24	17.6	15	2	A60834	angiotensin I prec
20	24	17.6	15	2	D48394	major fat-globule
21	24	17.6	17	2	S71864	glutathione transf
22	24	17.6	17	2	B31769	T-cell receptor de
23	24	17.6	18	2	PN0175	glutathione transf
24	24	17.6	19	2	B60822	cytochrome P450 UT
25	24	17.6	20	2	A37984	ADP,ATP carrier pr
26	24	17.6	21	2	A42762	multicatalytic end
27	24	17.6	22	2	A35418	brain natriuretic
28	23.5	17.3	26	2	S78374	hypothetical prote
29	23	16.9	14	2	PA0104	protein QF200070 -
30	23	16.9	16	2	A12729	glutamate dehydrog
31	23	16.9	19	2	PH1756	T cell receptor al
32	23	16.9	19	2	I46654	T-cell receptor de
33	23	16.9	20	2	A41439	acid ribonuclease
34	23	16.9	21	2	T07683	proteinase inhibit
35	23	16.9	24	2	T42257	phosphoprotein pho
36	23	16.9	24	2	A19828	female protein - h
37	23	16.9	25	2	I51565	protein-tyrosine k
38	23	16.9	26	2	A42218	early protein Sxl
39	22	16.2	7	1	A61324	dermorphin - Rohde
40	22	16.2	10	2	S13224	virG protein - Agr
41	22	16.2	11	2	PT0250	Ig heavy chain CRD
42	22	16.2	12	2	C36201	1-aminocyclopropan
43	22	16.2	14	2	A61002	photosystem II oxy
44	22	16.2	15	2	S71306	heat shock protein
45	22	16.2	17	2	C37520	glutathione transf

ALIGNMENTS

RESULT 1
B59018
MUC1 enhancer binding protein 85K chain MUC1EBP-85 - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
C;Accession: B59018
R;Abe, M.; Smith, C.J.; Larson, C.J.
submitted to the Protein Sequence Database, May 1998
A;Description: Involvement of "Ku-like" proteins in the transcription of MUC1/DF3, a
A;Reference number: A59018
A;Accession: B59018
A;Molecule type: protein
A;Residues: 1-17;18-26 <ABE>
A;Experimental source: breast cancer cell line MCF-7
C;Keywords: DNA binding; heterodimer

Query Match 21.3%; Score 29; DB 2; Length 26;
Best Local Similarity 37.5%; Pred. No. 3.7e+02;
Matches 6; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy       5 LLLQMDFGFPKHLLVD 20
           ::| || ||    |::
Db       9 VVLXMDVGFTLFPLIE 24
```

RESULT 2
A39504
octamer-binding protein, Ku-like, 72K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39504
R;May, G.; Sutton, C.; Gould, H.
J. Biol. Chem. 266, 3052-3059, 1991
A;Title: Purification and characterization of Ku-2, an octamer-binding protein relate
A;Reference number: A39504; MUID:91131605
A;Accession: A39504
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <MAY>

Query Match 20.6%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 3.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       7 LQMDFGF 13
             |:|| ||
Db      11 LEMDVGF 17
```

RESULT 3

```
;      FILING DATE:  07 June 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CRAIG, ANNE I.
;      REGISTRATION NUMBER:  32,976
;      REFERENCE/DOCKET NUMBER:  017.6US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 466-6000
;      TELEFAX:  (617) 466-6040
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-482-142-22


  Query Match             21.3%;  Score 29;  DB 2;  Length 26;
  Best Local Similarity   41.7%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 DGCLLLQMDFGF 13
           ||  ::| | |:
Db      11 DGTRIIQRDNGY 22


RESULT  15
US-08-478-572-22
; Sequence 22, Application US/08478572
; Patent No. 5968526
;  GENERAL INFORMATION:
;    APPLICANT:  Garman, Richard
;    APPLICANT:  Greenstein, Julia
;    APPLICANT:  Kuo, Mei-chang
;    APPLICANT:  Rogers, Bruce
;    APPLICANT:  Franzen, Henry
;    APPLICANT:  Chen, Xian
;    APPLICANT:  Evans, Sean
;    APPLICANT:  Shaked, Ze'ev
;    TITLE OF INVENTION:  T CELL EPITOPES OF THE MAJOR ALLERGENS
;    TITLE OF INVENTION:  FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
;    NUMBER OF SEQUENCES:  207
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  IMMULOGIC PHARMACEUTICAL CORPORATION
;      STREET:  610 LINCOLN STREET
;      CITY:  WALTHAM
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII TEXT
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,572
;      FILING DATE:  07-June-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/445,307
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CRAIG, ANNE I.
;      REGISTRATION NUMBER:  32,976
;      REFERENCE/DOCKET NUMBER:  017.6US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 466-6000
;      TELEFAX:  (617) 466-6040
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  N-terminal
US-08-478-572-22


  Query Match             21.3%;  Score 29;  DB 2;  Length 26;
  Best Local Similarity   41.7%;  Pred. No. 1.5e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 DGCLLLQMDFGF 13
           ||  ::| | |:
Db      11 DGTRIIQRDNGY 22


Search completed: December 21, 2000, 08:39:07
Job time: 141 sec
```

```
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111-4187
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/827,618A
;       FILING DATE:  09-APR-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/485,725
;       FILING DATE:  07-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/716,909
;       FILING DATE:  18-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/586,536
;       FILING DATE:  21-SEP-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Trecartin, Richard F.
;       REGISTRATION NUMBER:  31,801
;       REFERENCE/DOCKET NUMBER:  A-60780-12/RFT/MTK
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-827-618A-22


  Query Match             21.3%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPKHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT  13
US-08-483-952A-22
; Sequence 22, Application US/08483952A
; Patent No. 6011139
;  GENERAL INFORMATION:
;     APPLICANT:  Tobin, Allan J
;     APPLICANT:  Erlander, Mark G
;     APPLICANT:  Kaufman, Daniel L.
;     TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;     NUMBER OF SEQUENCES:  60
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;       STREET:  Four Embarcadero Center, Suite 3400
;       CITY:  San Francisco
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  94111-4187
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/483,952A
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/123,859
;       FILING DATE:  17-SEP-1993
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/716,909
;       FILING DATE:  18-JUN-1991
;       CLASSIFICATION:  536
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/586,536
;       FILING DATE:  21-SEP-1990
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Trecartin, Richard F.
;       REGISTRATION NUMBER:  31,801
;       REFERENCE/DOCKET NUMBER:  A-60780-7/RFT/MTK
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415-781-1989
;       TELEFAX:  415-398-3249
;   INFORMATION FOR SEQ ID NO:  22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-483-952A-22


  Query Match             21.3%;  Score 29;  DB 3;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPKHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT  14
US-08-482-142-22
; Sequence 22, Application US/08482142
; Patent No. 5820862
;  GENERAL INFORMATION:
;     APPLICANT:  Garman, Richard
;     APPLICANT:  Greenstein, Julia
;     APPLICANT:  Kuo, Mei-chang
;     APPLICANT:  Rogers, Bruce
;     APPLICANT:  Franzen, Henry
;     APPLICANT:  Chen, Xian
;     APPLICANT:  Evans, Sean
;     APPLICANT:  Shaked, Ze'ev
;     TITLE OF INVENTION:  T CELL EPITOPES OF THE MAJOR ALLERGENS
;     TITLE OF INVENTION:  FROM DERMATOPHAGOIDES (HOUSE DUST MITE)
;     NUMBER OF SEQUENCES:  207
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  IMMULOGIC PHARMACEUTICAL CORPORATION
;       STREET:  610 LINCOLN STREET
;       CITY:  WALTHAM
;       STATE:  MA
;       COUNTRY:  USA
;       ZIP:  02154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII TEXT
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/482,142
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/445,307
```

```
;            TELEPHONE: 919-248-1000
;   INFORMATION FOR SEQ ID NO: 142:
;        SEQUENCE CHARACTERISTICS:
;            LENGTH: 18 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: <Unknown>
;            TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-08-973-225-142


  Query Match             21.3%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   36.8%;  Pred. No. 99;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy       2 DGCLLLQMDFGFPKHLLVD 20
           ||| |        :| |:|
Db       3 DGCTL--------RHWLMD 13


RESULT  10
US-09-244-298A-142
; Sequence 142, Application US/09244298A
; Patent No. 6121238
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/244,298A
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  142:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-244-298A-142


  Query Match             21.3%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   36.8%;  Pred. No. 99;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy       2 DGCLLLQMDFGFPKHLLVD 20
           ||| |        :| |:|
Db       3 DGCTL--------RHWLMD 13


RESULT  11
US-08-484-530-22
; Sequence 22, Application US/08484530
; Patent No. 5846740
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisoc
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,530
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-8/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-530-22


  Query Match             21.3%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 MDFGFPKHLL 18
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT  12
US-08-827-618A-22
; Sequence 22, Application US/08827618A
; Patent No. 5998366
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
```

```
;    TITLE OF INVENTION:  HUMAN PLATELET GLYCOPROTEIN Ib/IX
;    NUMBER OF SEQUENCES:  173
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle LLP
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/556,597
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/406,330
;      FILING DATE:  17-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  20884/101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 263-1636
;      TELEFAX:  (716) 263-1600
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-556-597-60


  Query Match             21.3%;  Score 29;  DB 2;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 61;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LDGCLLL 7
            ||||  |
Db       2 LDGCFFL 8


RESULT   8
US-08-764-640-142
; Sequence 142, Application US/08764640
; Patent No. 5869451
; Patent No. 5869451 5837683
;  GENERAL INFORMATION:
;    APPLICANT:  Dower, William J.
;    APPLICANT:  Barrett, Ronald W.
;    APPLICANT:  Cwirla, Steven E.
;    APPLICANT:  Gates, Christian
;    APPLICANT:  Schatz, Peter J.
;    APPLICANT:  Balasubramanian, Palaniappan
;    APPLICANT:  Wagstrom, Christopher R.
;    APPLICANT:  Hendren, Richard W.
;    APPLICANT:  Deprince, Randolph B.
;    APPLICANT:  Podduturi, Surekha
;    APPLICANT:  Yin, Qun
;    TITLE OF INVENTION:  PEPTIDES AND COMPOUNDS THAT BIND TO A
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  244
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Glaxo Wellcome
;      STREET:  Five Moore Drive, P.O. Box 13398
;      CITY:  Research Triangle Park
;      STATE:  NC
;      COUNTRY:  USA
;      ZIP:  27709
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/764,640
;      FILING DATE:  11-DEC-1996
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hrubiec, Robert T.
;      REGISTRATION NUMBER:  36,392
;      REFERENCE/DOCKET NUMBER:  PK3281
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-248-1000
;  INFORMATION FOR SEQ ID NO:  142:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-764-640-142


  Query Match             21.3%;  Score 29;  DB 2;  Length 18;
  Best Local Similarity   36.8%;  Pred. No. 99;
  Matches    7;  Conservative    2;  Mismatches    2;  Indels    8;  Gaps    1;

Qy       2 DGCLLLQMDFGFPKHLLVD 20
            |||  |         :|  |:|
Db       3 DGCTL--------RHWLMD 13


RESULT   9
US-08-973-225-142
; Sequence 142, Application US/08973225A
; Patent No. 6083913
;   GENERAL INFORMATION:
;        APPLICANT: Dower, William J.
;                   Barrett, Ronald W.
;                   Cwirla, Steven E.
;                   Duffin, David J.
;                   Gates, Christian
;                   Haselden, Sherril S.
;                   Mattheakis, Larry C.
;                   Schatz, Peter J.
;                   Wagstrom, Christopher R.
;                   Wrighton, Nicholas C.
;        TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
;                            THROMBOPOIETIN RECEPTOR
;        NUMBER OF SEQUENCES: 232
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Glaxo Wellcome
;             STREET: Five Moore Drive, P.O. Box 13398
;             CITY: Research Triangle Park
;             STATE: NC
;             COUNTRY: USA
;             ZIP: 27709
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/973,225A
;             FILING DATE: 04-Dec-1997
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Hrubiec, Robert T.
;             REGISTRATION NUMBER: 36,392
;             REFERENCE/DOCKET NUMBER: PK3065USW
;        TELECOMMUNICATION INFORMATION:
```

```
;       CLASSIFICATION:  422
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Shannon, Karen L.
;       REGISTRATION NUMBER:  36,675
;       REFERENCE/DOCKET NUMBER:  8597/6
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-321-4200
;       TELEFAX:  312-321-4299
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  26 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-620-151-17


  Query Match             21.7%;  Score 29.5;  DB 2;  Length 26;
  Best Local Similarity   40.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    3;  Mismatches    5;  Indels    1;  Gaps    1;

Qy       4 CLLLQMDFGFPKHLL 18
           | :    ||: ||:|
Db       4 CKICSRSFGY-KHVL 17


RESULT   5
US-08-765-783A-83
; Sequence 83, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-83


  Query Match             21.3%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity   62.5%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 QMDFGFPK 15
           |  ||||:
Db       1 QHHFGFPR 8


RESULT   6
US-08-406-330-60
; Sequence 60, Application US/08406330
; Patent No. 5817748
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Jonathan L.
;    APPLICANT:  Lyle, Vicki A.
;    TITLE OF INVENTION:  MIMOTOPES AND ANTI-MIMOTOPES OF
;    TITLE OF INVENTION:  HUMAN PLATELET GLYCOPROTEIN Ib/IX
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nixon, Hargrave, Devans & Doyle
;      STREET:  Clinton Square, P.O. Box 1051
;      CITY:  Rochester
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  14603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/406,330
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Timian, Susan J.
;      REGISTRATION NUMBER:  34,103
;      REFERENCE/DOCKET NUMBER:  20884/100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (716) 263-1636
;      TELEFAX:  (716) 263-1600
;  INFORMATION FOR SEQ ID NO:  60:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  12 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-406-330-60


  Query Match             21.3%;  Score 29;  DB 2;  Length 12;
  Best Local Similarity   71.4%;  Pred. No. 61;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LDGCLLL 7
           ||||  |
Db       2 LDGCFFL 8


RESULT   7
US-08-556-597-60
; Sequence 60, Application US/08556597
; Patent No. 5877155
;  GENERAL INFORMATION:
;    APPLICANT:  Miller, Jonathan L.
;    APPLICANT:  Lyle, Vicki A.
;    TITLE OF INVENTION:  MIMOTOPES AND ANTI-MIMOTOPES OF
```

```
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    6;  Gaps    1;

Qy       3 GCLLLQMDFGFPKHLL 18
           |||      | | ||:
Db       2 GCL------GIPGHLM 11


RESULT   2
US-08-039-778B-2
; Sequence 2, Application US/08039778B
; Patent No. 5436137
;  GENERAL INFORMATION:
;    APPLICANT:  Spindel, Eliot R.
;    APPLICANT:  Vijayaraghavan, Srinivasan
;    APPLICANT:  Nagalla, Srinivasa R.
;    APPLICANT:  Li, Kang
;    TITLE OF INVENTION:  DNA SEQUENCE WHICH ENCODES A
;    TITLE OF INVENTION:  PEPTIDE CAPABLE OF PROMOTING
;    TITLE OF INVENTION:  ACROMSOME REACTION
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/039,778B
;      FILING DATE:  19930323
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/919,731
;      FILING DATE:  27 July, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00537/068002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-039-778B-2


  Query Match             22.8%;  Score 31;  DB 1;  Length 22;
  Best Local Similarity   43.8%;  Pred. No. 57;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    6;  Gaps    1;

Qy       3 GCLLLQMDFGFPKHLL 18
           |||      | | ||:
Db      13 GCL------GIPGHLM 22


RESULT   3
US-08-199-508-2
; Sequence 2, Application US/08199508
; Patent No. 5717058
;  GENERAL INFORMATION:
;    APPLICANT:  Matthews, Maura-Ann H.
;    APPLICANT:  Stetler, Gary L.
;    APPLICANT:  Anthony-Cahill, Spencer J.
;    APPLICANT:  Anderson, David C.
;    TITLE OF INVENTION:  Modulators of Gene Expression
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  5797 Central Avenue
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
;      COMPUTER:  Apple Macintosh
;      OPERATING SYSTEM:  System 7.0.1
;      SOFTWARE:  Microsoft Word 5.0a
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/199,508
;      FILING DATE:  February 18, 1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/021,536
;      FILING DATE:  February 23, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5717058ak, Henry P.
;      REGISTRATION NUMBER:  33200
;      REFERENCE/DOCKET NUMBER:  121 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  303-541-3322
;      TELEFAX:  303-444-3013
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  amino acid
;      TOPOLOGY:  unknown to applicant
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  yes
US-08-199-508-2


  Query Match             22.1%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 76;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      11 FGFPKHLLVDFLQ 23
           ||:| ::  | :|
Db       1 FGYPVYVFGDCVQ 13


RESULT   4
US-08-620-151-17
; Sequence 17, Application US/08620151
; Patent No. 5928955
;  GENERAL INFORMATION:
;    APPLICANT:  Imperiali, Barbara
;    APPLICANT:  Walkup, Grant K.
;    TITLE OF INVENTION:  PEPTIDYL FLUORESCENT CHEMOSENSOR FOR
;    TITLE OF INVENTION:  DIVALENT ZINC
;    NUMBER OF SEQUENCES:  136
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/620,151
;      FILING DATE:  22-MAR-1996
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:39:07 ; Search time 17.58 Seconds
(without alignments)
24.789 Million cell updates/sec

Title: US-08-934-367-50
Perfect score: 136
Sequence: 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 99815

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	22.8	11	1	US-08-039-778B-3	Sequence 3, Appli
2	31	22.8	22	1	US-08-039-778B-2	Sequence 2, Appli
3	30	22.1	20	1	US-08-199-508-2	Sequence 2, Appli
4	29.5	21.7	26	2	US-08-620-151-17	Sequence 17, Appl
5	29	21.3	9	2	US-08-765-783A-83	Sequence 83, Appl
6	29	21.3	12	2	US-08-406-330-60	Sequence 60, Appl
7	29	21.3	12	2	US-08-556-597-60	Sequence 60, Appl
8	29	21.3	18	2	US-08-764-640-142	Sequence 142, App
9	29	21.3	18	3	US-08-973-225-142	Sequence 142, App
10	29	21.3	18	3	US-09-244-298A-142	Sequence 142, App
11	29	21.3	20	2	US-08-484-530-22	Sequence 22, Appl
12	29	21.3	20	2	US-08-827-618A-22	Sequence 22, Appl
13	29	21.3	20	3	US-08-483-952A-22	Sequence 22, Appl
14	29	21.3	26	2	US-08-482-142-22	Sequence 22, Appl
15	29	21.3	26	2	US-08-478-572-22	Sequence 22, Appl
16	28	20.6	9	3	US-08-341-018-40	Sequence 40, Appl
17	28	20.6	17	1	US-08-370-567-7	Sequence 7, Appli
18	28	20.6	17	1	US-08-438-759-7	Sequence 7, Appli
19	28	20.6	17	4	PCT-US94-05684-7	Sequence 7, Appli
20	28	20.6	21	3	US-08-802-981-112	Sequence 112, App
21	28	20.6	21	3	US-08-802-981-113	Sequence 113, App
22	28	20.6	22	3	US-08-940-095-133	Sequence 133, App
23	28	20.6	22	3	US-08-940-093-133	Sequence 133, App
24	28	20.6	22	3	US-08-940-096-133	Sequence 133, App
25	28	20.6	26	1	US-08-475-989-27	Sequence 27, Appl
26	28	20.6	26	2	US-08-482-142-37	Sequence 37, Appl
27	28	20.6	26	2	US-08-482-142-134	Sequence 134, App
28	28	20.6	26	2	US-08-482-142-168	Sequence 168, App
29	28	20.6	26	2	US-08-478-572-37	Sequence 37, Appl
30	28	20.6	26	2	US-08-478-572-134	Sequence 134, App
31	28	20.6	26	2	US-08-478-572-168	Sequence 168, App
32	28	20.6	26	2	US-08-475-985-27	Sequence 27, Appl
33	28	20.6	26	3	US-08-256-839-27	Sequence 27, Appl
34	28	20.6	26	4	PCT-US95-04481-13	Sequence 13, Appl
35	28	20.6	26	4	PCT-US95-04481-30	Sequence 30, Appl
36	27.5	20.2	26	3	US-08-630-172-40	Sequence 40, Appl
37	27	19.9	9	2	US-08-482-651-39	Sequence 39, Appl
38	27	19.9	15	2	US-08-482-142-85	Sequence 85, Appl
39	27	19.9	15	2	US-08-478-572-85	Sequence 85, Appl
40	27	19.9	19	1	US-07-977-696C-36	Sequence 36, Appl
41	27	19.9	19	1	US-08-129-930B-36	Sequence 36, Appl
42	27	19.9	20	1	US-08-399-696-88	Sequence 88, Appl
43	27	19.9	20	2	US-08-733-505A-48	Sequence 48, Appl
44	27	19.9	20	2	US-08-934-915-7	Sequence 7, Appli
45	27	19.9	20	2	US-08-706-741B-83	Sequence 83, Appl

ALIGNMENTS

```
RESULT 1
US-08-039-778B-3
; Sequence 3, Application US/08039778B
; Patent No. 5436137
;  GENERAL INFORMATION:
;    APPLICANT:  Spindel, Eliot R.
;    APPLICANT:  Vijayaraghavan, Srinivasan
;    APPLICANT:  Nagalla, Srinivasa R.
;    APPLICANT:  Li, Kang
;    TITLE OF INVENTION:  DNA SEQUENCE WHICH ENCODES A
;    TITLE OF INVENTION:  PEPTIDE CAPABLE OF PROMOTING
;    TITLE OF INVENTION:  ACROMSOME REACTION
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  225 Franklin Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  U.S.A.
;      ZIP:  02110-2804
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      COMPUTER:  IBM PS/2 Model 50Z or 55SX
;      OPERATING SYSTEM:  MS-DOS (Version 5.0)
;      SOFTWARE:  WordPerfect (Version 5.1)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/039,778B
;      FILING DATE:  19930323
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/919,731
;      FILING DATE:  27 July, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clark, Paul T.
;      REGISTRATION NUMBER:  30,162
;      REFERENCE/DOCKET NUMBER:  00537/068002
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 542-5070
;      TELEFAX:  (617) 542-8906
;      TELEX:  200154
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  11
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
US-08-039-778B-3
```

Query Match 22.8%; Score 31; DB 1; Length 11;
Best Local Similarity 43.8%; Pred. No. 25;

OS Homo sapiens.
XX
PN WO9924836-A1.
XX
PD 20-MAY-1999.
XX
PF 04-NOV-1998; 98WO-US23435.
XX
PR 17-NOV-1997; 97US-0066100.
PR 07-NOV-1997; 97US-0064900.
PR 07-NOV-1997; 97US-0064908.
PR 07-NOV-1997; 97US-0064911.
PR 07-NOV-1997; 97US-0064912.
PR 07-NOV-1997; 97US-0064983.
PR 07-NOV-1997; 97US-0064984.
PR 07-NOV-1997; 97US-0064985.
PR 07-NOV-1997; 97US-0064987.
PR 07-NOV-1997; 97US-0064988.
PR 17-NOV-1997; 97US-0066090.
PR 17-NOV-1997; 97US-0066094.
PR 17-NOV-1997; 97US-0066095.
PR 17-NOV-1997; 97US-0066089.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;
PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y;
XX
DR WPI; 1999-337740/28.
DR N-PSDB; X85016.
XX
PT New human secreted proteins and coding sequences useful for treating
PT disorders of the immune system and hyperproliferative disorders
XX
PS Disclosure; Page 119; 507pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. X84924) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 125 novel genes and their fragments (nucleic
CC acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 125 polynucleotides, based on
CC which tissues they are most highly expressed in (see X84933 for described
CC uses).
XX
SQ Sequence 14 AA;

Query Match 23.5%; Score 32; DB 20; Length 14;
Best Local Similarity 61.5%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy      14 PKHLLVDFLQSLS 26
           |    | ||||||
Db       1 phssrvsflqsls 13
```

Search completed: December 21, 2000, 08:38:46
Job time: 150 sec

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /label= AA No. 40
XX
PN   WO8302461-A.
XX
PD   21-JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU-) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; N30158.
XX
PT   Multi:class hybrid interferon poly:peptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 17; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it.  Pref. the
CC   AA sequence consists of alpha and beta interferons.  Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
CC   N30160, P30227).  Alternativeley IF1 is the amino terminal end of a
CC   beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC   1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC   resp.) (see N30156, P30223).  In the examples plasmids pGW5 and
CC   pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,
CC   N30157).  HinfI was used to digest the DNA sequences in the region
CC   of significant handicaps (see N30153, N30154, N30158, N30159), and
CC   the restriction fragments were ligated to form hybrid DNA.
XX
SQ   Sequence   12 AA;

  Query Match             23.5%;  Score 32;  DB 4;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 48;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 DFGFPK 15
           |||||:
Db       2 dfgfpq 7


RESULT  14
P80053
ID   P80053 standard; protein; 12 AA.
XX
AC   P80053;
XX
DT   17-NOV-1990  (first entry)
XX
DE   Sequence of human interferon (huIFN) alpha-61A gene around AA 40.
XX
KW   Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW   antitumour; antiviral; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note="Residue 40"
XX
PN   US4758428-A.
XX
PD   19-JUL-1988.
XX
PF   15-JUL-1985;   85US-0755265.
XX
PR   15-JUL-1985;   85US-0755265.
PR   19-JAN-1983;   83CA-0419758.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1988-219882/31.
DR   N-PSDB; n80050.
XX
PT   Multi-class hybrid interferon polypeptide(s) -
PT   having sequence from interferon-alpha-1 and sequence from
PT   interferon-beta-1 for restricted activity
XX
PS   Example; Fig 17; 24pp; English.
XX
CC   Multi-class hybrid IFN polypeptides having an AA sequence composed
CC   of 2 distinct subsequences are claimed. The plasmids used in the
CC   construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC   pDM101/trp/beta-1. Assembly of the palpha61A plasmid invovled replacing
CC   the DNA fragment encoding the 23 AA signal polypeptide of preinterferon
CC   with a 120BP EcoRI/Sau3A promoter fragment E.coli trp promoter, operator,
CC   and trp leader ribosome binding site preceoperator, encoding an ATG
CC   initiation codon and using HindIII site that was inserted, 59 nucleotides
CC   3'-end of the TGA translational stop codon, to insert the gene into the
CC   plasmid pBW11 (a deriv of pBR322 having a deletion between the HindIII
CC   and PvuII sites). The complete DNA sequence of the promoter and gene
CC   fragments inserted between the EcoRI and HindIII sites of pBW11 is shown
CC   in n80049. The hybrid gene was constructed by taking advantage of the
CC   homologies between huIFN alpha-61A & huIFN beta-1 at around AA 40 of both
CC   proteins. The DNA sequence 5'-proximal to the DdeI restriction enzyme
CC   cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC   3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC   two genes while preserving the translational reading frame of both.
XX
SQ   Sequence   12 AA;

  Query Match             23.5%;  Score 32;  DB 9;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 48;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 DFGFPK 15
           |||||:
Db       2 dfgfpq 7


RESULT  15
Y27814
ID   Y27814 standard; Protein; 14 AA.
XX
AC   Y27814;
XX
DT   30-JUL-1999  (first entry)
XX
DE   Human secreted protein encoded by gene No. 84.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
```

```
Qy      11 FGFPKHLLVDFLQSLS 26
           ||||:||||||||||||
Db       1 fgfpehllvdflqsls 16


RESULT  11
W24294
ID   W24294 standard; peptide; 11 AA.
XX
AC   W24294;
XX
DT   17-OCT-1997  (first entry)
XX
DE   Human/Rabbit CETP common peptide.
XX
KW   Immune response; high density lipoprotein; HDL; cholesterol; human;
KW   serum; epitope; cholesteryl ester transfer protein; CETP; rabbit.
XX
OS   Oryctolagus cuniculus.
OS   Homo sapiens.
XX
PN   WO9639168-A1.
XX
PD   12-DEC-1996.
XX
PF   05-JUN-1996;   96WO-US09143.
XX
PR   06-JUN-1995;   95US-0482454.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW,  Carlo DJ,  Kwoh DY;
XX
DR   WPI; 1997-042849/04.
XX
PT   Stimulating an immune response to increase high density lipoprotein
PT   - avoids repeated administration of toxic drugs to lower cholesteryl
PT   ester transfer protein levels
XX
PS   Claim 5; Page 16; 26pp; English.
XX
CC   The sequences given in W24292-94 were used in the method of the
CC   invention to stimulate an immune response to increase high density
CC   lipoprotein (HDL) cholesterol in a mammal exhibiting low levels of
CC   serum HDL.  These peptides represent immunogenic epitopes of
CC   cholesteryl ester transfer protein (CETP).  The method utilises
CC   the body's own immune system to lower CETP levels, thereby increasing
CC   the level of beneficial HDL cholesterol, preferably in serum.  The
CC   method avoids the problems associated with the repeated administration
CC   of drugs which have undesirable side effects.  This peptide represents
CC   a region of CETP which is common to both human and rabbit proteins.
XX
SQ   Sequence   11 AA;


  Query Match             39.0%;  Score 53;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.017;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 HLLVDFLQSLS 26
           ||||||||||||
Db       1 hllvdflqsls 11


RESULT  12
W38080
ID   W38080 standard; Peptide; 21 AA.
XX
AC   W38080;
XX
DT   23-APR-1998  (first entry)
XX
DE   PPPPY motif containing peptide used to bind WW domains.
XX
KW   Peptide recognition unit; WW domain; cell signalling; growth regulation;
KW   cytoskeleton organisation; targeted drug screening; modulator;
KW   WW domain interaction; YAP protein; dystrophin.
XX
OS   Synthetic.
XX
PN   WO9737223-A1.
XX
PD   09-OCT-1997.
XX
PF   03-APR-1997;   97WO-US05547.
XX
PR   03-APR-1996;   96US-0630916.
XX
PA   (CYTO-) CYTOGEN CORP.
PA   (UYNC-) UNIV NORTH CAROLINA.
XX
PI   Fowlkes DM,  Kay BK,  Pirozzi G;
XX
DR   WPI; 1997-503234/46.
XX
PT   Identifying cell signalling and growth regulatory polypeptides by
PT   reaction with multivalent recognition complex - polypeptides are
PT   useful in targetted drug selection
XX
PS   Disclosure; Fig 15C; 220pp; English.
XX
CC   Peptides W38068-92 contain PPPPY-like motifs. The PPPY motif is
CC   found in the proline rich regions of WBP-1 and WBP-2 proteins. Peptides
CC   containing this residue have been shown to bind the YAP WW domain, but
CC   not the WW domain from dystrophin or to a panel of SH3 domains. Peptides
CC   W38068-92 were biotinylated and complexed with alkaline streptavidin,
CC   and used in a cross affinity mapping experiment. They were tested for
CC   their ability to bind to the 3 WW domains of WWP4 (W36797), which were
CC   expressed as glutathione-S-transferase expression proteins. The present
CC   peptide, derived from Prkacg (sic), binds to some of the WW domains
CC   of the novel protein. The WW domain is a small functional domain. Its
CC   name is derived from the observation that two tryptophan residues, 1 in
CC   the amino terminal portion of the WW domain and 1 in the carboxyl
CC   terminal portion, are conserved. Most proteins containing WW domains
CC   have a function involving cell signalling and growth regulation or the
CC   organisation of the cytoskeleton. Polypeptides containing a WW domain are
CC   identified by treating a multivalent recognition unit complex that has
CC   selective binding affinity for a WW domain, with many polypeptides and
CC   identifying those with selective affinity for the complex. Proteins
CC   containing WW domains are used for targeted drug screening, i.e. to
CC   identify potential modulators of specific WW domain interactions.
XX
SQ   Sequence   21 AA;


  Query Match             24.3%;  Score 33;  DB 18;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 62;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 GCLLLQMDFGFP 14
           | |: :|  |||
Db       1 gvliyemavgfp 12


RESULT  13
P30225
ID   P30225 standard; Protein; 12 AA.
XX
AC   P30225;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -alpha-61A around amino acid 40.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
```

CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 61.0%; Score 83; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FGFPKHLLVDFLQSLS 26
||||||||||||||||
Db 1 fgfpkhllvdflqsls 16

RESULT 10
Y91229
ID Y91229 standard; peptide; 16 AA.
XX
AC Y91229;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:107.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 50; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 58.1%; Score 79; DB 21; Length 16;
Best Local Similarity 93.8%; Pred. No. 1.6e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Disclosure; Page 75; 99pp; English.
XX
CC This sequence represents an immunogenic fragment of the rabbit
CC cholesteryl ester transferase protein (CETP).
CC The invention relates to recombinant DNA vaccines that contain DNA
CC encoding CETP, which can be used for producing antibodies to lessen the
CC transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC method can provide an autogenic immunological process for lessening the
CC transfer of cholesteryl esters from HDL particles and for increasing the
CC HDL cholesterol concentration of a mammal whose blood also contains
CC CETP. The method may be useful in treating human pro-atherogenic
CC dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC method can have an effect that lasts for months as compared to the
CC short-term effects of the small molecule drugs now available.
XX
SQ Sequence 22 AA;

Query Match 78.7%; Score 107; DB 20; Length 22;
Best Local Similarity 95.5%; Pred. No. 6.8e-11;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LLLQMDFGFPKHLLVDFLQSLS 26
||||||||||:||||||||||
Db 1 lllqmdfgfpehllvdflqsls 22

RESULT 8
Y13821
ID Y13821 standard; peptide; 22 AA.
XX
AC Y13821;
XX
DT 08-JUL-1999 (first entry)
XX
DE Human CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Homo sapiens.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Disclosure; Page 88; 99pp; English.
XX
CC This sequence represents an immunogenic fragment of the human
CC cholesteryl ester transferase protein (CETP).
CC The invention relates to recombinant DNA vaccines that contain DNA
CC encoding CETP, which can be used for producing antibodies to lessen the
CC transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC method can provide an autogenic immunological process for lessening the
CC transfer of cholesteryl esters from HDL particles and for increasing the
CC HDL cholesterol concentration of a mammal whose blood also contains
CC CETP. The method may be useful in treating human pro-atherogenic
CC dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC method can have an effect that lasts for months as compared to the
CC short-term effects of the small molecule drugs now available.
XX
SQ Sequence 22 AA;

Query Match 78.7%; Score 107; DB 20; Length 22;
Best Local Similarity 95.5%; Pred. No. 6.8e-11;
Matches 21; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 LLLQMDFGFPKHLLVDFLQSLS 26
||||||||||:||||||||||
Db 1 lllqmdfgfpehllvdflqsls 22

RESULT 9
Y91230
ID Y91230 standard; peptide; 16 AA.
XX
AC Y91230;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:108.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for

```
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
CC   Note: Sequence Y91227 is also designated SEQ ID NO:106 in the
CC   specification.
XX
SQ   Sequence   26 AA;


  Query Match             86.0%;  Score 117;  DB 21;  Length 26;
  Best Local Similarity   92.0%;  Pred. No. 2e-12;
  Matches   23;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DGCLLLQMDFGFPKHLLVDFLQSLS 26
           || |||||||||:|||||||||||
Db       2 dgflllqmdfgfpehllvdflqsls 26


RESULT   6
Y13809
ID   Y13809 standard; peptide; 22 AA.
XX
AC   Y13809;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Example 1; Page 73; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             81.6%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LLLQMDFGFPKHLLVDFLQSLS 26
           ||||||||||||||||||||||
Db       1 lllqmdfgfpkhllvdflqsls 22


RESULT   7
Y13815
ID   Y13815 standard; peptide; 22 AA.
XX
AC   Y13815;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
```

PF 01-MAY-1996; 96WO-US06147.
XX
PR 01-MAY-1995; 95US-0432483.
XX
PA (TCEL-) T CELL SCI INC.
XX
PI Rittershaus CW, Thomas LJ;
XX
DR WPI; 1996-506103/50.
XX
PT Cholesteryl ester transfer protein B cell epitope linked to T cell
PT epitope - used to generate vaccine to regulate CETP activity for
PT decreasing the risk of developing a cardiovascular disease e.g.
PT atherosclerosis
XX
PS Claim 5; Page 41; 72pp; English.
XX
CC A B-cell epitope (W06128) comprising the C-terminal 26 amino acids
CC of human liver mature cholesteryl ester transfer protein (CETP)
CC (see also W06127) is involved in a neutral lipid binding or a
CC transfer activity of CETP. It can be linked to a universal or
CC broad range immunogenic T-cell epitope, such as that found at amino
CC acids 830-843 of tetanus toxoid protein, to produce a synthetic
CC vaccine (see also W06129) that elicits an immune response against
CC endogenous CETP activity, thereby treating or preventing
CC cardiovascular disease, such as atherosclerosis. It may also be
CC incorporated into a multivalent vaccine (see also W06131)
CC including another CETP B-cell epitope.
XX
SQ Sequence 26 AA;

Query Match 86.0%; Score 117; DB 17; Length 26;
Best Local Similarity 92.0%; Pred. No. 2e-12;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DGCLLLQMDFGFPKHLLVDFLQSLS 26
|| |||||||||:|||||||||||
Db 2 dgflllqmdfgfpehllvdflqsls 26

RESULT 4
Y13801
ID Y13801 standard; peptide; 26 AA.
XX
AC Y13801;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Claim 15; Page 85; 99pp; English.
XX
CC This sequence represents an immunogenic fragment of the rabbit
CC cholesteryl ester transferase protein (CETP).
CC The invention relates to recombinant DNA vaccines that contain DNA
CC encoding CETP, which can be used for producing antibodies to lessen the
CC transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC method can provide an autogenic immunological process for lessening the
CC transfer of cholesteryl esters from HDL particles and for increasing the
CC HDL cholesterol concentration of a mammal whose blood also contains
CC CETP. The method may be useful in treating human pro-atherogenic
CC dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC method can have an effect that lasts for months as compared to the
CC short-term effects of the small molecule drugs now available.
XX
SQ Sequence 26 AA;

Query Match 86.0%; Score 117; DB 20; Length 26;
Best Local Similarity 92.0%; Pred. No. 2e-12;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DGCLLLQMDFGFPKHLLVDFLQSLS 26
|| |||||||||:|||||||||||
Db 2 dgflllqmdfgfpehllvdflqsls 26

RESULT 5
Y91228
ID Y91228 standard; peptide; 26 AA.
XX
AC Y91228;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:106.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 49; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport

CC This sequence represents an immunogenic fragment of the rabbit
CC cholesteryl ester transferase protein (CETP).
CC The invention relates to recombinant DNA vaccines that contain DNA
CC encoding CETP, which can be used for producing antibodies to lessen the
CC transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC method can provide an autogenic immunological process for lessening the
CC transfer of cholesteryl esters from HDL particles and for increasing the
CC HDL cholesterol concentration of a mammal whose blood also contains
CC CETP. The method may be useful in treating human pro-atherogenic
CC dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC method can have an effect that lasts for months as compared to the
CC short-term effects of the small molecule drugs now available.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 136; DB 20; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26
||||||||||||||||||||||||||
Db 1 ldgclllqmdfgfpkhllvdflqsls 26

RESULT 2
Y91231
ID Y91231 standard; peptide; 26 AA.
XX
AC Y91231;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:109.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 136; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26
||||||||||||||||||||||||||
Db 1 ldgclllqmdfgfpkhllvdflqsls 26

RESULT 3
W06128
ID W06128 standard; Peptide; 26 AA.
XX
AC W06128;
XX
DT 07-FEB-1997 (first entry)
XX
DE Human cholesteryl ester transfer protein C-terminal B-cell epitope.
XX
KW Cholesteryl ester transfer protein; CETP; antigen; vaccine;
KW cardiovascular disease; atherosclerosis; B-cell epitope.
XX
OS Homo sapiens.
XX
PN WO9634888-A1.
XX
PD 07-NOV-1996.
XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:38:46 ; Search time 26.24 Seconds
(without alignments)
33.881 Million cell updates/sec

Title: US-08-934-367-50
Perfect score: 136
Sequence: 1 LDGCLLLQMDFGFPKHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 143201

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	136	100.0	26	20	Y13802	Rabbit CETP immuno
2	136	100.0	26	21	Y91231	Human cholesteryl
3	117	86.0	26	17	W06128	Human cholesteryl
4	117	86.0	26	20	Y13801	Rabbit CETP immuno
5	117	86.0	26	21	Y91228	Human cholesteryl
6	111	81.6	22	20	Y13809	Rabbit CETP immuno
7	107	78.7	22	20	Y13815	Rabbit CETP immuno
8	107	78.7	22	20	Y13821	Human CETP immunog
9	83	61.0	16	21	Y91230	Human cholesteryl
10	79	58.1	16	21	Y91229	Human cholesteryl
11	53	39.0	11	18	W24294	Human/Rabbit CETP
12	33	24.3	21	18	W38080	PPPPY motif contai
13	32	23.5	12	4	P30225	Sequence of interf
14	32	23.5	12	9	P80053	Sequence of human
15	32	23.5	14	20	Y27814	Human secreted pro
16	32	23.5	21	19	W65690	Fibronectin bindin
17	31.5	23.2	16	17	R98414	Interferon-gamma r
18	31	22.8	11	15	R47611	Bombesin-like pept
19	31	22.8	21	19	W65666	Fibronectin bindin
20	31	22.8	22	15	R47610	Bombesin-like pept
21	31	22.8	22	19	W65669	Peptide #13. Synt
22	30.5	22.4	15	19	Y21149	Human bcl2 proto-o
23	30	22.1	20	15	R61276	Transactivating pr
24	30	22.1	20	19	W41190	Tax protein fragme
25	30	22.1	25	20	Y12067	Human 5' EST secre
26	29	21.3	12	18	W32681	Human platelet gly
27	29	21.3	12	19	W71811	Mimotope capable o
28	29	21.3	18	18	W36737	Thrombopoietin rec
29	29	21.3	18	18	W09586	Thrombopoietin rec
30	29	21.3	18	21	Y52473	RTS60-1 mutation-c
31	29	21.3	20	16	R72269	Glutamic acid deca
32	29	21.3	20	21	Y59570	GAD65 fragment, pe
33	29	21.3	22	17	R95899	Fragment #4 of 7-a
34	29	21.3	24	18	W22196	Endogenous TCR alp
35	29	21.3	24	20	Y13053	Human secreted pro
36	29	21.3	25	21	Y51627	S. cerevisiae RAD1
37	29	21.3	26	15	R51744	Der p I derived pe
38	29	21.3	26	19	W71924	Dermatophagoides D
39	29	21.3	26	20	Y50373	Dermatophagoides s
40	28	20.6	9	15	R46908	GGF segment D'. H
41	28	20.6	9	15	R55648	GGF segment D'. H
42	28	20.6	9	16	R67236	Bovine glial cell
43	28	20.6	9	17	W09354	Bovine glial growt
44	28	20.6	9	17	R96068	Human glial growth
45	28	20.6	9	17	R86700	Bovine glial growt

ALIGNMENTS

RESULT 1
Y13802
ID Y13802 standard; peptide; 26 AA.
XX
AC Y13802;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Claim 15; Page 94; 99pp; English.
XX

RT "A novel C-terminal domain in the thyroid hormone receptor selectively
RT mediates thyroid hormone inhibition.";
RL J. Biol. Chem. 269:32713-32716(1994).
SQ SEQUENCE 16 AA; 1920 MW; 79A831A4F3C8E22F CRC64;

Query Match 21.0%; Score 21; DB 4; Length 16;
Best Local Similarity 40.0%; Pred. No. 4.1e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      11 KLFLHLLDFQ 20
           || : : | |  
Db       4 KLLMKVTDLQ 13
```

Search completed: December 21, 2000, 08:35:36
Job time: 599 sec

```
DE   MAST CELL PROTEASE 5 (FRAGMENT).
GN   MCPT5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SVJ; TISSUE=LIVER;
RA   Lunderius C., Hellman L.;
RL   Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF119364; AAD43901.1; -.
KW   Protease.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2075 MW;  6AA33AA6B28E4B7D CRC64;


  Query Match             22.0%;  Score 22;  DB 11;  Length 19;
  Best Local Similarity   83.3%;  Pred. No. 3.3e+03;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     12 LFLHLL 17
           | ||||
Db      4 LTLHLL 9


RESULT  12
Q49448
ID   Q49448      PRELIMINARY;      PRT;    20 AA.
AC   Q49448;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   UNCERTAIN (FRAGMENT).
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC   Mycoplasmataceae; Mycoplasma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G-37;
RA   Peterson S.N., Hu P., Bott K.F., Hutchison C.A.;
RL   Submitted (SEP-1993) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G-37;
RA   Peterson S.N.;
RL   Thesis (1992), Microbiology and Immunology,
RL   University of North Carolina Medical School, USA.
DR   EMBL; U01744; AAD10557.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   20 AA;  2424 MW;  BF8CE3999592F014 CRC64;


  Query Match             22.0%;  Score 22;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 3.4e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      4 QASYSQKKLF 13
          :| |  | ||
Db      4 KAYYVMKTLF 13


RESULT  13
O99182
ID   O99182      PRELIMINARY;      PRT;     7 AA.
AC   O99182;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CYTOCHROME OXIDASE I (FRAGMENT).
GN   COI.
OS   Pterolebias zonatus.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC   Atherinomorpha; Cyprinodontiformes; Aplocheiloidei; Aplocheilidae;
OC   Pterolebias.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Murphy W.J., Thomerson J.E., Collier G.E.;
RT   "A molecular phylogeny of Neotropical aplocheiloid killifishes
RT   (Cyprinodontiformes, Rivulidae).";
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF002591; AAD01074.1; -.
KW   Mitochondrion.
FT   NON_TER      1      1
SQ   SEQUENCE   7 AA;  899 MW;  672721F6CB572030 CRC64;


  Query Match             21.0%;  Score 21;  DB 8;  Length 7;
  Best Local Similarity   66.7%;  Pred. No. 3e+05;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     12 LFLHLL 17
           |: |||
Db      2 LYQHLL 7


RESULT  14
Q44693
ID   Q44693      PRELIMINARY;      PRT;    10 AA.
AC   Q44693;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   AMYLOLIQUEFACIENS TRPE GENE (3' END) AND TRPD GENE (5' END)
DE   (FRAGMENT).
OS   Bacillus amyloliquefaciens.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85006754.
RA   Yoshimura K., Uemura J., Seki T., Oshima Y.;
RT   "Construction of a promoter-probe vector for a Bacillus subtilis host
RT   by using the trpD+ gene of Bacillus amyloliquefaciens.";
RL   J. Bacteriol. 159:905-912(1984).
DR   EMBL; K02661; AAB05353.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   10 AA;  1266 MW;  D5121FC729D5A416 CRC64;


  Query Match             21.0%;  Score 21;  DB 2;  Length 10;
  Best Local Similarity   33.3%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      8 SQKKLFLHL 16
          ||:    |:
Db      2 SQRSFLFHM 10


RESULT  15
Q9UD41
ID   Q9UD41      PRELIMINARY;      PRT;    16 AA.
AC   Q9UD41;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   BETA-ISOFORM THYROID HORMONE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95105146.
RA   Flynn T.R., Hollenberg A.N., Cohen O., Menke J.B., Usala S.J.,
RA   Tollin S., Hegarty M.K., Wondisford F.E.;
```

```
RESULT   7
Q56610
ID   Q56610      PRELIMINARY;      PRT;    18 AA.
AC   Q56610;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ACCA (FRAGMENT).
GN   ACCA.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6706;
RX   MEDLINE; 97074686.
RA   Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,
RA   Morris J.G.;
RT   "Cloning and characterization of dnaE, encoding the catalytic subunit
RT   of replicative DNA polymerase III, from Vibrio cholerae strain
RT   C6706.";
RL   Gene 175:281-283(1996).
DR   EMBL; U30472; AAC44579.1; -.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2153 MW;  18EBCDAD212842EF CRC64;


  Query Match             23.0%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     14 LHLLDFQ 20
          |: |||:
Db      3 LNFLDFE 9


RESULT   8
Q9R5Q6
ID   Q9R5Q6      PRELIMINARY;      PRT;    14 AA.
AC   Q9R5Q6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   2-HALOBENZOATE 1,2-DIOXYGENASE COMPONENT A BETA SUBUNIT.
OS   Burkholderia cepacia (Pseudomonas cepacia).
OC   Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC   Burkholderia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92104974.
RA   Fetzner S., Muller R., Lingens F.;
RT   "Purification and some properties of 2-halobenzoate 1,2-dioxygenase, a
RT   two-component enzyme system from Pseudomonas cepacia 2CBS.";
RL   J. Bacteriol. 174:279-290(1992).
SQ   SEQUENCE   14 AA;  1544 MW;  0BC0478DE855A33B CRC64;


  Query Match             22.0%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   28.6%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    5;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSQKKLFL 14
          |::::||     |:
Db      1 TSLESSYLDVVAFI 14


RESULT   9
Q9UQJ2
ID   Q9UQJ2      PRELIMINARY;      PRT;    15 AA.
AC   Q9UQJ2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PRORELAXIN H2 PRECURSOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Garibay-Tupas J.;
RT   "Characterization of the human relaxins H1 and H2 5'-flanking
RT   regions.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF104935; AAD21961.1; -.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1806 MW;  E9AFA622F3A86818 CRC64;


  Query Match             22.0%;  Score 22;  DB 4;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     13 FLHLL 17
          | |||
Db      6 FFHLL 10


RESULT  10
Q9RLU4
ID   Q9RLU4      PRELIMINARY;      PRT;    19 AA.
AC   Q9RLU4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NISG PROTEIN (FRAGMENT).
GN   NISG.
OS   Lactococcus lactis.
OG   Plasmid pLEB513.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Wahlstroem G., Takala T., Saris P.E.J.;
RT   "Evidence for a mosaic structure of the Tn5481 in Lactococcus lactis
RT   N8.";
RL   DNA Seq. 9:245-261(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Saris P.E.J.;
RT   "Characterization of the nisFEG operon of the nisin Z producing
RT   Lactococcus lactis subsp. lactis N8 strain.";
RL   DNA Seq. 9:263-274(1998).
DR   EMBL; AJ000993; CAA04441.1; -.
KW   Plasmid.
FT   NON_TER      1      1
SQ   SEQUENCE   19 AA;  2139 MW;  6A429DC80E673613 CRC64;


  Query Match             22.0%;  Score 22;  DB 2;  Length 19;
  Best Local Similarity   33.3%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSQKKL 12
          | :  : |:||:
Db      7 TIILVALSKKKI 18


RESULT  11
Q9R1F0
ID   Q9R1F0      PRELIMINARY;      PRT;    19 AA.
AC   Q9R1F0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
```

```
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93015942.
RA   Kojima K., Ogawa H.K., Seno N., Yamamoto K., Irimura T., Osawa T.,
RA   Matsumoto I.;
RT   "Carbohydrate-binding proteins in bovine kidney have consensus amino
RT   acid sequences of annexin family proteins.";
RL   J. Biol. Chem. 267:20536-20539(1992).
SQ   SEQUENCE   19 AA;  2294 MW;  F025B1E02C50C419 CRC64;


  Query Match             25.0%;  Score 25;  DB 6;  Length 19;
  Best Local Similarity   41.7%;  Pred. No. 9.9e+02;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       7 YSQKKLFLHLLD 18
           ||     ||: |
Db       6 YSXNXHLLHVFD 17


RESULT   3
Q46963
ID   Q46963      PRELIMINARY;      PRT;    15 AA.
AC   Q46963;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   SPOT 42 RNA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CP78;
RX   MEDLINE; 80074983.
RA   Sahagan B.G., Dahlberg J.E.;
RT   "A small, unstable RNA molecule of Escherichia coli: spot 42 RNA. I.
RT   Nucleotide sequence analysis.";
RL   J. Mol. Biol. 131:573-592(1979).
DR   EMBL; X01895; CAA25985.1; -.
SQ   SEQUENCE   15 AA;  1725 MW;  5A51DAF170EA661E CRC64;


  Query Match             24.0%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 LFLHLLDF 19
           | ||:: |
Db       7 LLLHVIGF 14


RESULT   4
Q9S8A4
ID   Q9S8A4      PRELIMINARY;      PRT;    16 AA.
AC   Q9S8A4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   18 KDA COELIAC SERUM-REACTIVE GLYCOPROTEIN (FRAGMENT).
OS   Secale cereale (Rye).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Secale.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96283789.
RA   Rocher A., Calero M., Soriano F., Mendez E.;
RT   "Identification of major rye secalins as coeliac immunoreactive
RT   proteins.";
RL   Biochim. Biophys. Acta 1295:13-22(1996).
SQ   SEQUENCE   16 AA;  1787 MW;  4659C4A6B2B28861 CRC64;


  Query Match             24.0%;  Score 24;  DB 10;  Length 16;
  Best Local Similarity   41.7%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       1 TTVQASYSQKKL 12
           ||    | | :|
Db       3 TTXSQGYGQXQL 14


RESULT   5
Q9R583
ID   Q9R583      PRELIMINARY;      PRT;    18 AA.
AC   Q9R583;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   TRANSFERRIN-BINDING PROTEIN 1 (FRAGMENT).
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93307625.
RA   Griffiths E., Stevenson P., Byfield P., Ala'Aldeen D.A.,
RA   Borriello S.P., Holland J., Parsons T., Williams P.;
RT   "Antigenic relationships of transferrin-binding proteins from
RT   Neisseria meningitidis, N. gonorrhoeae and Haemophilus influenzae:
RT   cross-reactivity of antibodies to NH2-terminal peptides.";
RL   FEMS Microbiol. Lett. 109:85-91(1993).
SQ   SEQUENCE   18 AA;  1999 MW;  FC0F1FB0A86C77C4 CRC64;


  Query Match             24.0%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   55.6%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSQKK 11
           |||  :|:|
Db       3 VQAGQAQEK 11


RESULT   6
Q9UQJ1
ID   Q9UQJ1      PRELIMINARY;      PRT;    11 AA.
AC   Q9UQJ1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PRORELAXIN H1 PRECURSOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Garibay-Tupas J.;
RT   "Characterization of the human H1 relaxin 5'-flanking region.";
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF104934; AAD21967.1; -.
FT   NON_TER      11      11
SQ   SEQUENCE   11 AA;  1416 MW;  40C86818E9C729C7 CRC64;


  Query Match             23.5%;  Score 23.5;  DB 4;  Length 11;
  Best Local Similarity   66.7%;  Pred. No. 1e+03;
  Matches    6;  Conservative    2;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      11 KLFL-HLLD 18
           :||| |||:
Db       3 RLFLFHLLE 11
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:34 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-5
Perfect score: 100
Sequence: 1 TTVQASYSQKKLFLHLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	25	25.0	18	3	Q02414	Q02414 aspergillus
2	25	25.0	19	6	Q9TRL8	Q9trl8 bos taurus
3	24	24.0	15	2	Q46963	Q46963 escherichia
4	24	24.0	16	10	Q9S8A4	Q9s8a4 secale cere
5	24	24.0	18	2	Q9R583	Q9r583 neisseria m
6	23.5	23.5	11	4	Q9UQJ1	Q9uqj1 homo sapien
7	23	23.0	18	2	Q56610	Q56610 vibrio chol
8	22	22.0	14	2	Q9R5Q6	Q9r5q6 burkholderi
9	22	22.0	15	4	Q9UQJ2	Q9uqj2 homo sapien
10	22	22.0	19	2	Q9RLU4	Q9rlu4 lactococcus
11	22	22.0	19	11	Q9R1F0	Q9r1f0 mus musculu
12	22	22.0	20	2	Q49448	Q49448 mycoplasma
13	21	21.0	7	8	O99182	O99182 pterolebias
14	21	21.0	10	2	Q44693	Q44693 bacillus am
15	21	21.0	16	4	Q9UD41	Q9ud41 homo sapien
16	21	21.0	18	2	O06514	O06514 escherichia
17	21	21.0	18	11	P97522	P97522 rattus norv
18	21	21.0	19	2	O87732	O87732 streptomyce
19	21	21.0	19	3	Q92321	Q92321 saccharomyc
20	21	21.0	19	12	Q86583	Q86583 murine hepa
21	21	21.0	20	2	Q9R4W5	Q9r4w5 helicobacte
22	21	21.0	20	9	Q37868	Q37868 bacteriopha
23	21	21.0	20	12	Q86935	Q86935 human herpe
24	21	21.0	20	12	Q86936	Q86936 human herpe
25	21	21.0	20	12	Q86937	Q86937 human herpe
26	21	21.0	20	12	Q86938	Q86938 human herpe
27	21	21.0	20	12	Q86939	Q86939 human herpe
28	21	21.0	20	12	Q86940	Q86940 human herpe
29	21	21.0	20	12	Q86941	Q86941 human herpe
30	21	21.0	20	12	Q86942	Q86942 human herpe
31	20.5	20.5	14	2	Q9R2Z4	Q9r2z4 campylobact
32	20	20.0	7	12	Q9YQ10	Q9yq10 porcine tra
33	20	20.0	9	2	Q45852	Q45852 clostridium
34	20	20.0	14	5	O18502	O18502 schistosoma
35	20	20.0	15	2	Q9R542	Q9r542 mycobacteri
36	20	20.0	15	4	Q9UCH4	Q9uch4 homo sapien
37	20	20.0	16	2	Q47605	Q47605 escherichia
38	20	20.0	16	2	Q9R557	Q9r557 bacillus sp
39	20	20.0	16	12	Q79458	Q79458 human immun
40	20	20.0	18	2	P97166	P97166 streptococc
41	20	20.0	18	4	Q9UJ81	Q9uj81 homo sapien
42	20	20.0	18	13	Q91380	Q91380 gallus gall
43	20	20.0	19	4	Q07603	Q07603 homo sapien
44	20	20.0	19	4	Q9UMM9	Q9umm9 homo sapien
45	20	20.0	19	10	Q9S901	Q9s901 vigna sinen

ALIGNMENTS

RESULT 1
Q02414
ID Q02414 PRELIMINARY; PRT; 18 AA.
AC Q02414;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE OROTIDINE-5'-MONOPHOSPHATE DECARBOXYLASE (FRAGMENT).
GN PYRG.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
OC anamorphic Trichocomaceae; Aspergillus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96031709.
RA Gouka R.J., Hessing J.G., Stam H., Musters W., Hondel C.A.;
RT "A novel strategy for the isolation of defined pyrG mutants and the
RT development of a site-specific integration system for Aspergillus
RT awamori.";
RL Curr. Genet. 27:536-540(1995).
DR EMBL; S79674; AAB35350.1; -.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 1891 MW; E1A0E34C7D23688E CRC64;

Query Match 25.0%; Score 25; DB 3; Length 18;
Best Local Similarity 41.2%; Pred. No. 9.4e+02;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

```
Qy       2 TVQASYSQKKLFLHLLD 18
           || |  :  |  | | |
Db       2 TVSADVTTTKELLDLAD 18
```

RESULT 2
Q9TRL8
ID Q9TRL8 PRELIMINARY; PRT; 19 AA.
AC Q9TRL8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE CA(2+)-DEPENDENT CARBOHYDRATE-BINDING PROTEIN (FRAGMENT).

```
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X64731; CAA45996.1; -.
DR   PIR; S23423; S23423.
DR   MENDEL; 2272; PAVlu;atpI;1.
DR   INTERPRO; IPR000568; -.
DR   PROSITE; PS00449; ATPASE_A; PARTIAL.
KW   Hydrogen ion transport; CF(0); Chloroplast; Transmembrane.
FT   NON_TER       1      1
FT   TRANSMEM     <1     13       POTENTIAL.
SQ   SEQUENCE   17 AA;  1753 MW;  F313DCE74C23CF2E CRC64;


  Query Match             18.0%;  Score 18;  DB 1;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 3.7e+03;
  Matches    3;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TVQASY 7
          |: |:|
Db      6 TLSAAY 11


RESULT  14
SP51_BACLI
ID   SP51_BACLI     STANDARD;      PRT;    17 AA.
AC   P27642;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   STAGE V SPORULATION PROTEIN AA (FRAGMENT).
GN   SPOVAA.
OS   Bacillus licheniformis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91237317.
RA   Moldover B., Piggot P.J., Yudkin M.D.;
RT   "Identification of the promoter and the transcriptional start site of
RT   the spoVA operon of Bacillus subtilis and Bacillus licheniformis.";
RL   J. Gen. Microbiol. 137:527-531(1991).
CC   ---------------------------------------------------------------------------
CC | This SWISS-PROT entry is copyright. It is produced through a collaboration
CC | between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X53991; CAA37938.1; -.
DR   PIR; S16144; S16144.
KW   Sporulation.
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  2167 MW;  17505DA08BFE133B CRC64;


  Query Match             18.0%;  Score 18;  DB 1;  Length 17;
  Best Local Similarity   25.0%;  Pred. No. 3.7e+03;
  Matches    2;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 QKKLFLHL 16
          ::::|: |
Db      2 ERQVFIRL 9


RESULT  15
AL22_HORSE
ID   AL22_HORSE     STANDARD;      PRT;    19 AA.
AC   P81217;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   DANDER MAJOR ALLERGEN EQU C 2.0102 (FRAGMENT).
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 98237590.
RA   Bulone V., Krogstad-Johnsen T., Smestad-Paulsen B.;
RT   "Separation of horse dander allergen proteins by two-dimensional
RT   electrophoresis -- molecular characterisation and identification of
RT   Equ c 2.0101 and Equ c 2.0102 as lipocalin proteins.";
RL   Eur. J. Biochem. 253:202-211(1998).
CC   -!- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
CC   -!- PTM: N-GLYCOSYLATED.
CC   -!- MISCELLANEOUS: PREDOMINANT ALLERGEN OF HORSE DANDER.
CC   -!- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC   -!- CAUTION: C 2.0101 AND C 2.0102 MIGHT BE TWO VARIANTS OF THE
CC       SAME PROTEIN.
DR   INTERPRO; IPR000566; -.
DR   PROSITE; PS00213; LIPOCALIN; PARTIAL.
KW   Transport; Lipocalin; Glycoprotein; Allergen.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2195 MW;  A8EE6FAFC9322C26 CRC64;


  Query Match             18.0%;  Score 18;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 QASYSQ 9
          :  |||
Db      7 ETDYSQ 12


Search completed: December 21, 2000, 08:32:55
Job time: 438 sec
```

```
          ||||
Db      8 SQKK 11


RESULT  10
AMP_FUSNU
ID   AMP_FUSNU      STANDARD;      PRT;    20 AA.
AC   P81207;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   AMINOPEPTIDASE (EC 3.4.11.-) (AP) (FRAGMENT).
OS   Fusobacterium nucleatum.
OC   Bacteria; Fusobacteria; Fusobacterium.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=ATCC 10953;
RX   MEDLINE; 98361026.
RA   Rogers A.H., Gunadi A., Gully N.J., Zilm P.S.;
RT   "An aminopeptidase nutritionally important to Fusobacterium
RT   nucleatum.";
RL   Microbiology 144:1807-1813(1998).
CC   -!- FUNCTION: CLEAVES A WIDE RANGE OF DIPEPTIDES AND TRIPEPTIDES.
CC   -!- COFACTOR: COBALT.
CC   -!- SUBCELLULAR LOCATION: CELL ENVELOPE-ASSOCIATED.
KW   Hydrolase; Aminopeptidase; Metalloprotease; Cobalt.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2585 MW;  97095B948262C71B CRC64;


  Query Match             19.0%;  Score 19;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      10 KKLFLHLLDF 19
           |: ||  : |
Db      10 KERFLRYVKF 19


RESULT  11
DCMS_PSECF
ID   DCMS_PSECF     STANDARD;      PRT;    20 AA.
AC   P19915;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE SMALL CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0145; PL0145.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2155 MW;  9C7BD676A6C558AA CRC64;


  Query Match             19.0%;  Score 19;  DB 1;  Length 20;
  Best Local Similarity   40.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TVQASYSQKK 11


          ||    :|:|
Db      7 TVNVGKAQEK 16


RESULT  12
FIBA_HORSE
ID   FIBA_HORSE     STANDARD;      PRT;    14 AA.
AC   P14452;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-JAN-1990 (Rel. 13, Last annotation update)
DE   FIBRINOPEPTIDE A.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RA   Blomback B., Blomback M., Grondahl N.J.;
RT   "Studies on fibrinopeptides from mammals.";
RL   Acta Chem. Scand. 19:1789-1791(1965).
CC   -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC       POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC       AGGREGATION.
CC   -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC       (ALPHA, BETA, & GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC   -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC       THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC       CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC       RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
KW   Blood coagulation; Plasma.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1517 MW;  4E998EB63C2A15E7 CRC64;


  Query Match             18.0%;  Score 18;  DB 1;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 3e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 FLH 15
           |||
Db       6 FLH 8


RESULT  13
ATPI_PAVLU
ID   ATPI_PAVLU     STANDARD;      PRT;    17 AA.
AC   P28529;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   ATP SYNTHASE A CHAIN (EC 3.6.1.34) (SUBUNIT IV) (FRAGMENT).
GN   ATPI.
OS   Pavlova lutherii (Monochrysis lutheri).
OG   Chloroplast.
OC   Eukaryota; Haptophyceae; Pavlovales; Pavlova.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 92316212.
RA   Scaramuzzi C.D., Stokes H.W., Hiller R.G.;
RT   "Characterisation of a chloroplast-encoded secY homologue and atpH
RT   from a chromophytic alga. Evidence for a novel chloroplast genome
RT   organisation.";
RL   FEBS Lett. 304:119-123(1992).
CC   -!- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
CC       DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE.
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CHLOROPLAST
CC       THYLAKOID MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC   ---------------------------------------------------------------------
```

```
CC       TETRAHYDROCANNABINOL (THC) AND STRYCHNINE.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A36154; A36154.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2259 MW;  78DC81280C970A55 CRC64;


  Query Match             20.0%;  Score 20;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 2e+03;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 LFLHLL 17
           ||| ||
Db       7 LFLALL 12


RESULT   7
NUHM_CANFA
ID   NUHM_CANFA     STANDARD;      PRT;    11 AA.
AC   P49820;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   NADH-UBIQUINONE OXIDOREDUCTASE 24 KDA SUBUNIT (EC 1.6.5.3)
DE   (EC 1.6.99.3) (FRAGMENT).
GN   NDUFV2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC       CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC       TO BE UBIQUINONE. COMPONENT OF THE FLAVOPROTEIN-SULFUR (FP)
CC       FRAGMENT OF THE ENZYME.
CC   -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC   -!- COFACTOR: 2FE-2S IRON-SULFUR CLUSTER N 1B.
CC   -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 40 DIFFERENT SUBUNITS.
CC       THIS IS A COMPONENT OF THE FLAVOPROTEIN FRACTION.
CC   -!- SUBCELLULAR LOCATION: MATRIX AND CYTOPLASMIC SIDE OF THE
CC       MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE COMPLEX I 24 KDA SUBUNIT FAMILY.
DR   HSC-2DPAGE; P49820; DOG.
DR   INTERPRO; IPR002023; -.
DR   PROSITE; PS01099; COMPLEX1_24K; PARTIAL.
KW   Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Iron-sulfur.
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1099 MW;  267F5369C9C72DD8 CRC64;


  Query Match             19.0%;  Score 19;  DB 1;  Length 11;
  Best Local Similarity   75.0%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 LFLH 15
           ||:|
Db       6 LFVH 9


RESULT   8
PA21_MICFM
ID   PA21_MICFM     STANDARD;      PRT;    12 AA.
AC   P25072;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE   2-ACYLHYDROLASE) (FRAGMENT).
OS   Micrurus fulvius microgalbineus (Mexican coral snake).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Elapinae; Micrurus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 79255521.
RA   Possani L.D., Alagon A.C., Fletcher P.L. Jr., Varela M.J., Julia J.Z.;
RT   "Purification and characterization of a phospholipase A2 from the
RT   venom of the coral snake, Micrurus fulvius microgalbineus (Brown and
RT   Smith).";
RL   Biochem. J. 179:603-606(1979).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   INTERPRO; IPR001211; -.
DR   PROSITE; PS00118; PA2_HIS; PARTIAL.
DR   PROSITE; PS00119; PA2_ASP; PARTIAL.
KW   Hydrolase; Lipid degradation; Calcium; Venom.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1398 MW;  CC21992A899F0339 CRC64;


  Query Match             19.0%;  Score 19;  DB 1;  Length 12;
  Best Local Similarity   60.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 LLDFQ 20
           ||:|:
Db       2 LLBFK 6


RESULT   9
TX3_PHONI
ID   TX3_PHONI      STANDARD;      PRT;    19 AA.
AC   P31010;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   NEUROTOXIN TX3 (FRAGMENT).
OS   Phoneutria nigriventer (Brazilian armed spider).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC   Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phoneutria.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 92196803.
RA   Rezende L. Jr., Cordeiro M.N., Oliveira E.B., Diniz C.R.;
RT   "Isolation of neurotoxic peptides from the venom of the 'armed'
RT   spider Phoneutria nigriventer.";
RL   Toxicon 29:1225-1233(1991).
DR   PIR; C39305; C39305.
KW   Venom; Neurotoxin.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2244 MW;  3214E89CF10F7587 CRC64;


  Query Match             19.0%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 SQKK 11
```

```
RA   Passerini de Rossi B.N., Friedman L.E., Gonzalez Flecha F.L.,
RA   Castello P.R., Franco M.A., Rossi J.P.F.C.;
RT   "Identification of Bordetella pertussis virulence-associated outer
RT   membrane proteins.";
RL   FEMS Microbiol. Lett. 172:9-13(1999).
KW   Outer membrane; Virulence.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2116 MW;  D24E1CDCA665206C CRC64;


  Query Match             22.0%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 TVQASYS 8
          ||:  ||
Db     11 TVEGEYS 17


RESULT   3
MAST_POLJA
ID   MAST_POLJA     STANDARD;      PRT;     14 AA.
AC   P01517;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   POLISTES MASTOPARAN.
OS   Polistes jadwigae (Paper wasp).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Polistinae; Polistes.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Hirai Y., Ueno Y., Yasuhara T., Yoshida H., Nakajima T.;
RT   "A new mast cell degranulating peptide, polistes mastoparan, in the
RT   venom of Polistes jadwigae.";
RL   Biomed. Res. 1:185-187(1980).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC       THAT COUPLE TO PHOSPHOLIPASE C.
DR   PIR; A01780; QMWAPP.
KW   Mast cell degranulation; Venom; Amidation.
FT   MOD_RES      14     14       AMIDATION.
SQ   SEQUENCE   14 AA;  1636 MW;  26472A53BF4778D8 CRC64;


  Query Match             21.0%;  Score 21;  DB 1;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 9e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     10 KKLFLHLL 17
          ||:  |:|
Db      4 KKIGQHIL 11


RESULT   4
TL18_SPIOL
ID   TL18_SPIOL     STANDARD;      PRT;     20 AA.
AC   P82536;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   THYLAKOID LUMENAL 18 KDA PROTEIN (P18) (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RA   Kieselbach T., Bystedt M., Schroeder W.P.;
RL   Submitted (MAY-2000) to the SWISS-PROT data bank.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.
KW   Chloroplast; Thylakoid membrane.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2192 MW;  752C21963F49FA64 CRC64;


  Query Match             21.0%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   30.8%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSQKKLF 13
            | :|:  : | :|
Db      5 TPLQSKVTNKVVF 17


RESULT   5
XYNB_DICB4
ID   XYNB_DICB4     STANDARD;      PRT;     10 AA.
AC   P80717;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ENDO-1,4-BETA-XYLANASE B (EC 3.2.1.8) (XYLANASE B)
DE   (1,4-BETA-D-XYLAN XYLANOHYDROLASE B) (FRAGMENT).
OS   Dictyoglomus sp. (strain B4A).
OC   Bacteria; Dictyoglomus group; Dictyoglomus.
RN   [1]
RP   SEQUENCE.
RA   Adamsen A.K., Jacobsen S., Ahring B.K.;
RL   Submitted (OCT-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC       LINKAGES IN XYLANS.
CC   -!- PATHWAY: XYLAN DEGRADATION.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC       HYDROLASES).
DR   INTERPRO; IPR001000; -.
DR   PROSITE; PS00591; GLYCOSYL_HYDROL_F10; PARTIAL.
KW   Xylan degradation; Hydrolase; Glycosidase.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1144 MW;  4554322AA72041A3 CRC64;


  Query Match             20.0%;  Score 20;  DB 1;  Length 10;
  Best Local Similarity   55.6%;  Pred. No. 9.3e+02;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     10 KKLFLHLLD 18
          ||  | | |
Db      2 KKTILDLKD 10


RESULT   6
CPBX_CAVPO
ID   CPBX_CAVPO     STANDARD;      PRT;     20 AA.
AC   P34033;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CYTOCHROME P450IIB (EC 1.14.14.1) (FRAGMENT).
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN   [1]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   STRAIN=HARTLEY; TISSUE=LIVER;
RX   MEDLINE; 91054472.
RA   Narimatsu S., Akutsu Y., Matsunaga T., Watanabe K., Yamamoto I.,
RA   Yoshimura H.;
RT   "Purification of a cytochrome P450 isozyme belonging to a subfamily
RT   of P450 IIB from liver microsomes of guinea pigs.";
RL   Biochem. Biophys. Res. Commun. 172:607-613(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME IS ACTIVE
CC       UPON P.NITROANISOLE, ANILINE, D-BENZPHETAMINE, DELTA(9)-
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:54 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-5
Perfect score: 100
Sequence: 1 TTVQASYSQKKLFLHLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	24.0	18	1	UC21_MAIZE	P80627 zea mays (m
2	22	22.0	20	1	VR90_BORPE	P81549 bordetella
3	21	21.0	14	1	MAST_POLJA	P01517 polistes ja
4	21	21.0	20	1	TL18_SPIOL	P82536 spinacia ol
5	20	20.0	10	1	XYNB_DICB4	P80717 dictyoglomu
6	20	20.0	20	1	CPBX_CAVPO	P34033 cavia porce
7	19	19.0	11	1	NUHM_CANFA	P49820 canis famil
8	19	19.0	12	1	PA21_MICFM	P25072 micrurus fu
9	19	19.0	19	1	TX3_PHONI	P31010 phoneutria
10	19	19.0	20	1	AMP_FUSNU	P81207 fusobacteri
11	19	19.0	20	1	DCMS_PSECF	P19915 pseudomonas
12	18	18.0	14	1	FIBA_HORSE	P14452 equus cabal
13	18	18.0	17	1	ATPI_PAVLU	P28529 pavlova lut
14	18	18.0	17	1	SP51_BACLI	P27642 bacillus li
15	18	18.0	19	1	AL22_HORSE	P81217 equus cabal
16	18	18.0	19	1	UP24_UPEIN	P82030 uperoleia i
17	18	18.0	20	1	CRP_MUSCA	P19094 mustelus ca
18	17	17.0	9	1	HUTU_KLEAE	P12381 klebsiella
19	17	17.0	9	1	ULAE_HUMAN	P31931 homo sapien
20	17	17.0	10	1	PNEU_RAT	P21996 rattus norv
21	17	17.0	12	1	NO40_SOYBN	P55960 glycine max
22	17	17.0	15	1	CBPB_PROAT	P19628 protopterus
23	17	17.0	18	1	AL13_CARMA	P81816 carcinus ma
24	17	17.0	18	1	NODD_RHILO	Q52838 rhizobium l
25	17	17.0	20	1	CATA_ACIRA	P81422 acinetobact
26	17	17.0	20	1	RIPX_CUCPE	P80750 cucurbita p
27	16	16.0	8	1	AKH_MELML	P25423 melolontha
28	16	16.0	8	1	ANG2_BOTJA	Q10582 bothrops ja
29	16	16.0	8	1	COXG_RAT	P80430 rattus norv
30	16	16.0	10	1	TKNB_RANRI	P29135 rana ridibu
31	16	16.0	11	1	ANGT_CRIGE	P09037 crinia geor
32	16	16.0	12	1	PA2B_VIPBO	P31859 vipera beru
33	16	16.0	13	1	MLA_CAMDR	P01198 camelus dro
34	16	16.0	14	1	ANGT_HORSE	P01016 equus cabal
35	16	16.0	14	1	LECB_PSOSC	P22584 psophocarpu
36	16	16.0	14	1	RS19_CLOPP	Q46228 clover prol
37	16	16.0	14	1	RS19_LOWBP	Q48878 loofah witc
38	16	16.0	14	1	TKNM_RANMA	P40951 rana margar
39	16	16.0	15	1	COXJ_THUOB	P80979 thunnus obe
40	16	16.0	15	1	FGF1_CANFA	P18651 canis famil
41	16	16.0	15	1	PGKH_PHYPA	P80659 physcomitre
42	16	16.0	16	1	IBP4_PIG	P24854 sus scrofa
43	16	16.0	16	1	RIPK_TRIKI	P16093 trichosanth
44	16	16.0	17	1	UP41_UPEIN	P82035 uperoleia i
45	16	16.0	18	1	PHPT_PSESE	P25271 pseudaletia

ALIGNMENTS

RESULT 1
UC21_MAIZE
ID UC21_MAIZE STANDARD; PRT; 18 AA.
AC P80627;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 443)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE.
RC TISSUE=COLEOPTILE;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.9, ITS MW IS: 26.5 KDA.
CC -!- SIMILARITY: TO L-ASCORBATE PEROXIDASES.
DR HSSP; P48534; 1APX.
DR MAIZE-2DPAGE; P80627; COLEOPTILE.
DR MAIZEDB; 123953; -.
FT NON_TER 1 1
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1938 MW; F32F6FEF038BAB8A CRC64;

Query Match 24.0%; Score 24; DB 1; Length 18;
Best Local Similarity 71.4%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 TVQASYS 8
          || | ||
Db      6 TVSAEYS 12
```

RESULT 2
VR90_BORPE
ID VR90_BORPE STANDARD; PRT; 20 AA.
AC P81549;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE VIRULENCE-ASSOCIATED OUTER MEMBRANE PROTEIN VIR90 (FRAGMENT).
GN VIR90.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC Bordetella.
RN [1]
RP SEQUENCE.
RC STRAIN=TOHAMA I;
RX MEDLINE; 99179239.

A;Cross-references: GDB:119297; OMIM:141900
A;Map position: 11p15.4-11p15.4

Query Match 22.0%; Score 22; DB 4; Length 14;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VQASYSQ 9
|||:| :
Db 3 VQAAYQK 9

RESULT 13
B45895
T-cell surface glycoprotein CD28 short form - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Feb-1994
C;Accession: B45895
R;Lee, K.P.; Taylor, C.; Petryniak, B.; Turka, L.A.; June, C.H.; Thompson, C.B.
J. Immunol. 145, 344-352, 1990
A;Title: The genomic organization of the CD28 gene. Implications for the regulation of C
A;Reference number: A45895; MUID:90293482
A;Accession: B45895
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16 <LEE>
A;Cross-references: GB:M37813
C;Keywords: glycoprotein

Query Match 22.0%; Score 22; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 1.9e+03;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VQASYSQK 10
| ||::|
Db 2 VNLSYNEK 9

RESULT 14
S00495
hemocyanin chain III - Japanese spiny lobster (fragment)
C;Species: Panulirus japonicus (Japanese spiny lobster)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 01-Aug-1997
C;Accession: S00495
R;Makino, N.; Kimura, S.
Eur. J. Biochem. 173, 423-430, 1988
A;Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.
A;Reference number: S00492; MUID:88196131
A;Accession: S00495
A;Molecule type: protein
A;Residues: 1-19 <MAK>
C;Superfamily: hemocyanin
C;Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 22.0%; Score 22; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 2.3e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 HLLD 18
||||
Db 16 HLLD 19

RESULT 15
S00493
hemocyanin chain Ib - Japanese spiny lobster (fragment)
C;Species: Panulirus japonicus (Japanese spiny lobster)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 01-Aug-1997
C;Accession: S00493
R;Makino, N.; Kimura, S.
Eur. J. Biochem. 173, 423-430, 1988
A;Title: Subunits of Panulirus japonicus hemocyanin. 1. Isolation and properties.
A;Reference number: S00492; MUID:88196131
A;Accession: S00493
A;Molecule type: protein
A;Residues: 1-20 <MAK>
C;Superfamily: hemocyanin
C;Keywords: copper; hemolymph; hexamer; oxygen carrier

Query Match 22.0%; Score 22; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.4e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 15 HLLD 18
||||
Db 16 HLLD 19

Search completed: December 21, 2000, 08:30:04
Job time: 268 sec

C;Species: Aspergillus niger
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Sep-1997
C;Accession: S10452
R;van Gorcom, R.F.M.; Boschloo, J.G.; Kuijvenhoven, A.; Lange, J.; Bos, C.J.; van Balken
submitted to the EMBL Data Library, March 1990
A;Reference number: S10452
A;Accession: S10452
A;Molecule type: DNA
A;Residues: 1-18 <VAN>
A;Cross-references: EMBL:X52521; NID:g2336; PID:g2337

Query Match 23.0%; Score 23; DB 2; Length 18;
Best Local Similarity 42.9%; Pred. No. 1.4e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

```
Qy      2 TVQASYSQKKLFLH 15
          | |    |: || |
Db      2 TDQPLSRQETLFTH 15
```

RESULT 8
S54272
CTC 75 protein - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S54272
R;Genersch, E.; Eckerskorn, C.; Lottspeich, F.; Herzog, C.; Kuehn, K.; Poeschl, E.
EMBO J. 14, 791-800, 1995
A;Title: Purification of the sequence-specific transcription factor CTCBF, involved in t
A;Reference number: S54272; MUID:95188883
A;Accession: S54272
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <GEN>

Query Match 23.0%; Score 23; DB 2; Length 18;
Best Local Similarity 33.3%; Pred. No. 1.4e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSQKKL 12
          :| :  ||:::|
Db      6 STGKVEYSEEEL 17
```

RESULT 9
PH0786
T-cell receptor alpha chain (I7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0786
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0786
A;Molecule type: mRNA
A;Residues: 1-13 <CAS>
A;Cross-references: EMBL:X60889
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 22.0%; Score 22; DB 2; Length 13;
Best Local Similarity 33.3%; Pred. No. 1.6e+03;
Matches 3; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```
Qy      5 ASYSQKKLF 13
          ::|: | :|
Db      4 SAYANKMIF 12
```

RESULT 10
S29209
avenin alpha-2 - oat (fragment)
N;Alternate names: CIP-3; coeliac immunoreactive protein 3
C;Species: Avena sativa (oat)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C;Accession: S29209
R;Rocher, A.; Colilla, F.; Ortiz, M.L.; Mendez, E.
FEBS Lett. 310, 37-40, 1992
A;Title: Identification of the three major coeliac immunoreactive proteins and one al
A;Reference number: S29207; MUID:92405739
A;Accession: S29209
A;Molecule type: protein
A;Residues: 1-14 <ROC>
A;Experimental source: endosperm
C;Superfamily: gliadin
C;Keywords: prolamin; seed

Query Match 22.0%; Score 22; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 1.7e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSQK 10
          |||| : |::
Db      2 TTVQYNPSEQ 11
```

RESULT 11
A44920
2-halobenzoate 1,2-dioxygenase component A beta chain - Pseudomonas cepacia (fragment
C;Species: Pseudomonas cepacia
C;Date: 01-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C;Accession: A44920
R;Fetzner, S.; Muller, R.; Lingens, F.
J. Bacteriol. 174, 279-290, 1992
A;Title: Purification and some properties of 2-halobenzoate 1,2-dioxygenase, a two-co
A;Reference number: A44920; MUID:92104974
A;Contents: 2CBS
A;Accession: A44920
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <FET>
A;Note: sequence extracted from NCBI backbone (NCBIP:75379)

Query Match 22.0%; Score 22; DB 2; Length 14;
Best Local Similarity 28.6%; Pred. No. 1.7e+03;
Matches 4; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

```
Qy      1 TTVQASYSQKKLFL 14
          |:::::||     |:
Db      1 TSLESSYLDVVAFI 14
```

RESULT 12
I52618
hemoglobin beta chain thalassemia mutant Portuguese - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 11-Jul-1996 #text_change 20-Apr-2000
C;Accession: I52618
R;Oner, R.; Oner, C.; Wilson, J.B.; Tamagnini, G.P.; Ribeiro, L.M.; Huisman, T.H.
Br. J. Haematol. 79, 306-310, 1991
A;Title: Dominant beta-thalassaemia trait in a Portuguese family is caused by a delet
A;Reference number: I52618; MUID:92068764
A;Accession: I52618
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <ONE>
A;Cross-references: GB:S68042; NID:g239717; PIDN:AAB20440.1; PID:g239718
C;Genetics:
A;Gene: GDB:HBB

RESULT 3
S14560
probable heme-binding protein - garden pea chloroplast (fragment)
C;Species: chloroplast Pisum sativum (garden pea)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C;Accession: S14560
R;Smith, A.G.; Wilson, R.J.; Kaethner, T.M.; Willey, D.L.; Gray, J.C.
submitted to the EMBL Data Library, October 1990
A;Reference number: S14557
A;Accession: S14560
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <SMI>
A;Cross-references: EMBL:X54750; NID:g12194; PID:g12198
C;Genetics:
A;Genome: chloroplast
C;Superfamily: maize chloroplast protein cemA
C;Keywords: chloroplast; heme; transmembrane protein

Query Match 25.0%; Score 25; DB 2; Length 18;
Best Local Similarity 50.0%; Pred. No. 6.6e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 8 SQKKLFLHLL 17
::|| |: ||
Db 2 AKKKAFIPLL 11

RESULT 4
A26228
spot 42 protein - Escherichia coli
C;Species: Escherichia coli
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 08-Oct-1999
C;Accession: A26228; A37586; Q00551
R;Joyce, C.M.; Grindley, N.D.F.
J. Bacteriol. 152, 1211-1219, 1982
A;Title: Identification of two genes immediately downstream from the polA gene of Escher
A;Reference number: A26228; MUID:83056713
A;Accession: A26228
A;Molecule type: DNA
A;Residues: 1-15 <JOY2>
R;Sahagan, B.G.; Dahlberg, J.E.
J. Mol. Biol. 131, 573-592, 1979
A;Title: A small, unstable RNA molecule of Escherichia coli: spot 42 RNA.
A;Reference number: A37586; MUID:80074983
A;Accession: A37586
A;Molecule type: DNA
A;Residues: 1-15 <SAH>
A;Cross-references: GB:X01895; NID:g40868; PIDN:CAA25985.1; PID:g40869
C;Genetics:
A;Gene: spf
A;Map position: 87 min

Query Match 24.0%; Score 24; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 8.1e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 12 LFLHLLDF 19
| ||:: |
Db 7 LLLHVIGF 14

RESULT 5
S70331
endosperm protein, 18K - rye (fragment)
C;Species: Secale cereale (rye)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S70331
R;Rocher, A.; Calero, M.; Soriano, F.; Mendez, E.
Biochim. Biophys. Acta 1295, 13-22, 1996
A;Title: Identification of major rye secalins as coeliac immunoreactive proteins.
A;Reference number: S70327; MUID:96283789
A;Accession: S70331
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-16 <ROC>

Query Match 24.0%; Score 24; DB 2; Length 16;
Best Local Similarity 41.7%; Pred. No. 8.7e+02;
Matches 5; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TTVQASYSQKKL 12
|| | | :|
Db 3 TTXSQGYGQXQL 14

RESULT 6
S47365
T-cell antigen receptor VJ junction beta chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47365; S47375; S47379; S47396; S47397; S47398; S47355
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by
A;Reference number: S47355
A;Accession: S47365
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LEH>
A;Cross-references: EMBL:Z35690; NID:g527471; PIDN:CAA84759.1; PID:g527472; EMBL:Z356
A;Accession: S47375
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LE2>
A;Cross-references: EMBL:Z35700; NID:g527493; PIDN:CAA84769.1; PID:g527494
A;Accession: S47379
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LE3>
A;Cross-references: EMBL:Z35708; NID:g527509; PIDN:CAA84777.1; PID:g527510
A;Accession: S47396
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LE4>
A;Cross-references: EMBL:Z35674; NID:g527527; PIDN:CAA84743.1; PID:g527528
A;Accession: S47397
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LE5>
A;Cross-references: EMBL:Z35675; NID:g527529; PIDN:CAA84744.1; PID:g527530
A;Accession: S47398
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-13 <LE6>
A;Cross-references: EMBL:Z35676; NID:g527531; PIDN:CAA84745.1; PID:g527532
C;Keywords: T-cell receptor

Query Match 23.0%; Score 23; DB 2; Length 13;
Best Local Similarity 33.3%; Pred. No. 1e+03;
Matches 3; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTVQASYSQ 9
:::::|| |
Db 3 SSIRSSYEQ 11

RESULT 7
S10452
hypothetical protein (bphA 5' region) - Aspergillus niger

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:03 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-5
Perfect score: 100
Sequence: 1 TTVQASYSQKKLFLHLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	28.0	14	2	S13864	methyl coenzyme M
2	26	26.0	15	2	S65717	prostaglandin D-sy
3	25	25.0	18	2	S14560	probable heme-bind
4	24	24.0	15	2	A26228	spot 42 protein -
5	24	24.0	16	2	S70331	endosperm protein,
6	23	23.0	13	2	S47365	T-cell antigen rec
7	23	23.0	18	2	S10452	hypothetical prote
8	23	23.0	18	2	S54272	CTC 75 protein - h
9	22	22.0	13	2	PH0786	T-cell receptor al
10	22	22.0	14	2	S29209	avenin alpha-2 - o
11	22	22.0	14	2	A44920	2-halobenzoate 1,2
12	22	22.0	14	4	I52618	hemoglobin beta ch
13	22	22.0	16	2	B45895	T-cell surface gly
14	22	22.0	19	2	S00495	hemocyanin chain I
15	22	22.0	20	2	S00493	hemocyanin chain I
16	22	22.0	20	2	S00494	hemocyanin chain I
17	21	21.0	6	2	A61411	ameletin - rat
18	21	21.0	10	2	I40032	trpE protein - Bac
19	21	21.0	12	2	S65136	kallikrein K2 - hu
20	21	21.0	13	2	S47374	T-cell antigen rec
21	21	21.0	13	2	PH1479	T-cell receptor be
22	21	21.0	14	1	QMWAPP	polistes mastopara
23	21	21.0	16	2	A36889	leu operon leader
24	21	21.0	16	2	S34444	blaZ protein - Sta
25	21	21.0	17	2	JQ2320	hypothetical 2.1K
26	21	21.0	19	2	A48354	nonstructural prot
27	21	21.0	19	2	S59717	hypothetical prote
28	21	21.0	20	2	H49034	nuclear antigen EB
29	20	20.0	12	2	S26556	T-cell receptor be
30	20	20.0	13	2	S47372	T-cell antigen rec
31	20	20.0	15	2	PN0164	hyoscyamine (6S)-d
32	20	20.0	15	2	S36890	ribosomal protein
33	20	20.0	15	2	PH0770	T-cell receptor be
34	20	20.0	16	2	PC1299	subtilisin (EC 3.4
35	20	20.0	17	2	B49404	T-cell receptor be
36	20	20.0	17	2	I49593	cystic fibrosis tr
37	20	20.0	17	2	I84733	gene CFTR protein
38	20	20.0	17	2	A35550	adrenocortical cel
39	20	20.0	19	2	A60459	aminopeptidase, 30
40	20	20.0	19	2	I39327	pre-T/NK cell-asso
41	20	20.0	20	2	A60372	pollen allergen Po
42	20	20.0	20	2	S46479	retinoid-X-recepto
43	19	19.0	13	2	A44818	extracellular lipa
44	19	19.0	14	2	A47421	leukotriene B-4 12
45	19	19.0	15	2	PA0046	protein QA100044 -

ALIGNMENTS

RESULT 1
S13864
methyl coenzyme M reductase (EC 1.8.-.-) II alpha chain - Methanobacterium thermoauto
C;Species: Methanobacterium thermoautotrophicum
A;Variety: strain Marburg
C;Date: 19-Mar-1997 #sequence_revision 13-Sep-1998 #text_change 30-Oct-1998
C;Accession: S13864
R;Rospert, S.; Linder, D.; Ellermann, J.; Thauer, R.K.
Eur. J. Biochem. 194, 871-877, 1990
A;Title: Two genetically distinct methyl-coenzyme M reductases in Methanobacterium th
A;Reference number: S13864; MUID:91099370
A;Accession: S13864
A;Molecule type: protein
A;Residues: 1-14 <ROS>
A;Experimental source: strain Marburg
C;Keywords: methanogenesis; oxidoreductase

Query Match 28.0%; Score 28; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      9 QKKLFLHLL 17
          :|||||  |
Db      3 EKKLFLKAL 11
```

RESULT 2
S65717
prostaglandin D-synthase - rat (fragment)
N;Alternate names: prostaglandin-H2 D-isomerase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 31-Oct-1997
C;Accession: S65717
R;Giacomelli, S.; Leone, M.G.; Grima, J.; Silvestrini, B.; Cheng, C.Y.
Biochim. Biophys. Acta 1310, 269-276, 1996
A;Title: Astrocytes synthesize and secrete prostaglandin D synthetase in vitro.
A;Reference number: S65716; MUID:96177373
A;Accession: S65717
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <GIA>
C;Superfamily: lipocalin; lipocalin homology

Query Match 26.0%; Score 26; DB 2; Length 15;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 TVQASYSQKK 11
          ||| :: | |
Db      3 TVQPNFQQDK 12
```

```
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-383-753-57


  Query Match             27.0%;  Score 27;  DB 1;  Length 18;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 SQKKLFLHL 16
           :|| |: ||
Db      10 AQKMLYQHL 18


RESULT  14
US-08-586-772-57
; Sequence 57, Application US/08586772
; Patent No. 5874239
;  GENERAL INFORMATION:
;    APPLICANT:  Schatz, Peter J.
;    TITLE OF INVENTION:  Biotinylation of Proteins
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Steuart Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/586,772
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,991
;      FILING DATE:  30-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  1038.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-586-772-57


  Query Match             27.0%;  Score 27;  DB 2;  Length 18;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 SQKKLFLHL 16
           :|| |: ||
Db      10 AQKMLYQHL 18


RESULT  15
US-08-959-512-57
; Sequence 57, Application US/08959512
; Patent No. 5932433
;  GENERAL INFORMATION:
;    APPLICANT:  Schatz, Peter J.
;    TITLE OF INVENTION:  Biotinylation of Proteins
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Steuart Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/959,512
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/383,753
;      FILING DATE:  03-FEB-1995
;      APPLICATION NUMBER:  US 08/099,991
;      FILING DATE:  30-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  1038.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-959-512-57


  Query Match             27.0%;  Score 27;  DB 2;  Length 18;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       8 SQKKLFLHL 16
           :|| |: ||
Db      10 AQKMLYQHL 18


Search completed: December 21, 2000, 08:31:46
Job time: 369 sec
```

```
  Best Local Similarity   83.3%;  Pred. No. 80;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      15 HLLDFQ 20
           :|||||
Db       2 YLLDFQ 7


RESULT  11
US-09-100-414B-2
; Sequence 2, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-2


  Query Match             27.0%;  Score 27;  DB 3;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 1e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 KKLFLHLLD 18
           | | || ||
Db       6 KGLLLHKLD 14


RESULT  12
US-09-100-409A-61
; Sequence 61, Application US/09100409A
; Patent No. 6090388
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  PEPTIDE COMPOSITION FOR
;    TITLE OF INVENTION:  PREVENTION AND TREATMENT OF HIV INFECTION AND
;    TITLE OF INVENTION:  IMMUNE DISORDERS
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version
;      SOFTWARE:  #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,409A
;      FILING DATE:
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:
;      REGISTRATION NUMBER:
;      REFERENCE/DOCKET NUMBER:  1151-4154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-409A-61


  Query Match             27.0%;  Score 27;  DB 3;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 1e+02;
  Matches    6;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 KKLFLHLLD 18
           | | || ||
Db       6 KGLLLHKLD 14


RESULT  13
US-08-383-753-57
; Sequence 57, Application US/08383753
; Patent No. 5723584
;  GENERAL INFORMATION:
;    APPLICANT:  Schatz, Peter J.
;    TITLE OF INVENTION:  Biotinylation of Proteins
;    NUMBER OF SEQUENCES:  102
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  One Market Plaza, Steuart Tower
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/383,753
;      FILING DATE:  03-FEB-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/099,991
;      FILING DATE:  30-JUL-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M.
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  1038.1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-326-2400
;      TELEFAX:  415-326-2422
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/854,029
;      FILING DATE:  2 MAY 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Vincent, Matthew P.
;      REGISTRATION NUMBER:  36,709
;      REFERENCE/DOCKET NUMBER:  MIV-019.08
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-832-1242
;      TELEFAX:  617-832-7000
;  INFORMATION FOR SEQ ID NO:  28:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-854-029-28


  Query Match             28.0%;  Score 28;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 58;
  Matches    5;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 YSQKKLF 13
            |||::||
Db       8 YSQEELF 14


RESULT   8
5266328-11
;Patent No. 5266328
;    APPLICANT: SKUBITZ, AMY P.N.;FURCHT, LEO T.
;    TITLE OF INVENTION: LAMININ  CHAIN POLYPEPTIDES FROM
;THE CARBOXY TERMINAL GLOBULAR DOMAIN
;    NUMBER OF SEQUENCES: 14
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/573,672
;      FILING DATE: 27-AUG-1990
;SEQ ID NO:11:
;      LENGTH: 17
5266328-11


  Query Match             28.0%;  Score 28;  DB 5;  Length 17;
  Best Local Similarity   41.7%;  Pred. No. 73;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TVQASYSQKKLF 13
            ||:  | ::| |
Db       6 TVKTEYIKRKAF 17


RESULT   9
US-07-654-839-5
; Sequence 5, Application US/07654839
; Patent No. 5372933
;  GENERAL INFORMATION:
;    APPLICANT:  Zamarron, Concepcion
;    APPLICANT:  Plow, Edward F
;    APPLICANT:  Ginsberg, Mark H
;    TITLE OF INVENTION:  MONOCLONAL ANTIBODIES AGAINST
;    TITLE OF INVENTION:  RECEPTOR-INDUCED BINDING SITES
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10666 No. 5372933th Torrey Pines Road., Suite 220,
;      STREET:  Mail Drop TPC8
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/654,839
;      FILING DATE:  19910213
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/252,753
;      FILING DATE:  03-OCT-1988
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/415,029
;      FILING DATE:  29-SEP-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bingham, Douglas A
;      REGISTRATION NUMBER:  32,457
;      REFERENCE/DOCKET NUMBER:  SCR0367P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-554-2937
;      TELEFAX:  619-554-6312
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-07-654-839-5


  Query Match             28.0%;  Score 28;  DB 1;  Length 19;
  Best Local Similarity   38.5%;  Pred. No. 83;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 YSQKKLFLHLLDF 19
            || ||  : :: |
Db       1 YSMKKTTMKIIPF 13


RESULT  10
PCT-US94-01234-47
; Sequence 47, Application PC/TUS9401234
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  METHODS FOR PRODUCING POLYPEPTIDE
;    TITLE OF INVENTION:  BINDING SITES
;    NUMBER OF SEQUENCES:  76
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/01234
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/084,542
;      FILING DATE:  28-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/012,566
;      FILING DATE:  02-FEB-1993
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
PCT-US94-01234-47


  Query Match             27.0%;  Score 27;  DB 4;  Length 13;
```

; FILING DATE: 18-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: CCOR-0183
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-454-207A-43

Query Match 29.0%; Score 29; DB 1; Length 9;
Best Local Similarity 85.7%; Pred. No. 1.2e+05;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 8 SQKKLFL 14
|:|||||
Db 1 SKKKLFL 7

RESULT 5
US-08-428-415-28
; Sequence 28, Application US/08428415
; Patent No. 5756335
; GENERAL INFORMATION:
; APPLICANT: Cold Spring Harbor Laboratory
; TITLE OF INVENTION: No. 5756335el Human cdc25 Genes, Encoded
; TITLE OF INVENTION: Products and Uses Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,415
; FILING DATE: 24 April 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-019CP3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-428-415-28

Query Match 28.0%; Score 28; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 YSQKKLF 13
|||::||
Db 8 YSQEELF 14

RESULT 6
US-08-379-685-28
; Sequence 28, Application US/08379685
; Patent No. 5770423
; GENERAL INFORMATION:
; APPLICANT: Cold Spring Harbor Laboratory
; TITLE OF INVENTION: No. 5770423el Human cdc25 Genes, Encoded
; TITLE OF INVENTION: Products and Uses Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/379,685
; FILING DATE: 26 January 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MII-019-DV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-227-7400
; TELEFAX: 617-227-5941
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-379-685-28

Query Match 28.0%; Score 28; DB 1; Length 14;
Best Local Similarity 71.4%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 YSQKKLF 13
|||::||
Db 8 YSQEELF 14

RESULT 7
US-08-854-029-28
; Sequence 28, Application US/08854029
; Patent No. 5994074
; GENERAL INFORMATION:
; APPLICANT: Cold Spring Harbor Laboratory
; TITLE OF INVENTION: No. 5994074el Human cdc25 Genes, Encoded
; TITLE OF INVENTION: Products and Uses Thereof
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)

```
Qy      11 KLFLHLL 17
           ||| |||
Db       6 KLFFHLL 12


RESULT   2
US-08-871-355A-355
; Sequence 355, Application US/08871355A
; Patent No. 6015669
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/871,355A
;      FILING DATE:  09-JUN-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101 CON
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  355:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-871-355A-355


  Query Match             31.0%;  Score 31;  DB 3;  Length 14;
  Best Local Similarity   85.7%;  Pred. No. 17;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 KLFLHLL 17
           ||| |||
Db       6 KLFFHLL 12


RESULT   3
US-08-454-207A-36
; Sequence 36, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-36


  Query Match             29.0%;  Score 29;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 SQKKLFLH 15
            |:||| ||
Db       1 SKKKLALH 8


RESULT   4
US-08-454-207A-43
; Sequence 43, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:45 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-5
Perfect score: 100
Sequence: 1 TTVQASYSQKKLFLHLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	31.0	14	2	US-08-637-759B-355	Sequence 355, App
2	31	31.0	14	3	US-08-871-355A-355	Sequence 355, App
3	29	29.0	9	1	US-08-454-207A-36	Sequence 36, Appl
4	29	29.0	9	1	US-08-454-207A-43	Sequence 43, Appl
5	28	28.0	14	1	US-08-428-415-28	Sequence 28, Appl
6	28	28.0	14	1	US-08-379-685-28	Sequence 28, Appl
7	28	28.0	14	2	US-08-854-029-28	Sequence 28, Appl
8	28	28.0	17	5	5266328-11	Patent No. 5266328
9	28	28.0	19	1	US-07-654-839-5	Sequence 5, Appli
10	27	27.0	13	4	PCT-US94-01234-47	Sequence 47, Appl
11	27	27.0	16	3	US-09-100-414B-2	Sequence 2, Appli
12	27	27.0	16	3	US-09-100-409A-61	Sequence 61, Appl
13	27	27.0	18	1	US-08-383-753-57	Sequence 57, Appl
14	27	27.0	18	2	US-08-586-772-57	Sequence 57, Appl
15	27	27.0	18	2	US-08-959-512-57	Sequence 57, Appl
16	26	26.0	19	1	US-08-267-092A-49	Sequence 49, Appl
17	26	26.0	19	2	US-08-540-412-50	Sequence 50, Appl
18	26	26.0	19	4	PCT-US95-08156-50	Sequence 50, Appl
19	26	26.0	20	1	US-08-218-025A-81	Sequence 81, Appl
20	25	25.0	6	2	US-08-482-228-208	Sequence 208, App
21	25	25.0	6	3	US-08-482-528-208	Sequence 208, App
22	25	25.0	10	2	US-08-482-228-209	Sequence 209, App
23	25	25.0	10	3	US-08-482-528-209	Sequence 209, App
24	25	25.0	10	3	US-08-946-329A-42	Sequence 42, Appl
25	25	25.0	11	1	US-08-464-531-53	Sequence 53, Appl
26	25	25.0	11	2	US-08-461-598-53	Sequence 53, Appl
27	25	25.0	11	3	US-08-322-137-53	Sequence 53, Appl
28	25	25.0	14	2	US-08-844-031-23	Sequence 23, Appl
29	25	25.0	15	2	US-08-726-464B-35	Sequence 35, Appl
30	25	25.0	17	3	US-09-192-048-16	Sequence 16, Appl
31	25	25.0	18	3	US-09-100-414B-26	Sequence 26, Appl
32	25	25.0	18	3	US-09-100-414B-30	Sequence 30, Appl
33	25	25.0	20	1	US-07-678-974D-16	Sequence 16, Appl
34	25	25.0	20	2	US-08-934-915-179	Sequence 179, App
35	25	25.0	20	2	US-08-945-168-21	Sequence 21, Appl
36	24	24.0	10	2	US-08-967-999-15	Sequence 15, Appl
37	24	24.0	11	2	US-08-310-912A-117	Sequence 117, App
38	24	24.0	11	4	PCT-US95-04589-117	Sequence 117, App
39	24	24.0	16	1	US-08-158-232-44	Sequence 44, Appl
40	24	24.0	16	2	US-08-611-928-44	Sequence 44, Appl
41	24	24.0	16	3	US-09-100-437-1	Sequence 1, Appli
42	24	24.0	16	3	US-09-173-891-44	Sequence 44, Appl
43	24	24.0	17	1	US-08-518-474-9	Sequence 9, Appli
44	24	24.0	17	1	US-08-518-474-11	Sequence 11, Appl
45	24	24.0	17	2	US-08-248-839C-144	Sequence 144, App

ALIGNMENTS

```
RESULT   1
US-08-637-759B-355
; Sequence 355, Application US/08637759B
; Patent No. 5876931
;  GENERAL INFORMATION:
;    APPLICANT:  David William Holden
;    TITLE OF INVENTION:  Identification of Genes
;    NUMBER OF SEQUENCES:  501
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patrea L. Pabst
;      STREET:  2800 One Atlantic Center
;      STREET:  1201 West Peachtree Street
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30309-3450
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/637,759B
;      FILING DATE:  03-MAY-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB95/02875
;      FILING DATE:  11-DEC-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Pabst, Patrea L.
;      REGISTRATION NUMBER:  31,284
;      REFERENCE/DOCKET NUMBER:  RPMS 101
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 873-8794
;      TELEFAX:  (404) 873-8795
;  INFORMATION FOR SEQ ID NO:  355:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-637-759B-355
```

Query Match 31.0%; Score 31; DB 2; Length 14;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE; 20196006.
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE002927; AAF45430.1; -.
DR FLYBASE; FBgn0040023; CG12575.
FT NON_TER 19 19
SQ SEQUENCE 19 AA; 2039 MW; FAF4913F8360228A CRC64;

Query Match 21.6%; Score 24; DB 5; Length 19;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy      12 HLLVD 16
           ||:||
Db       4 HLIVD 8
```

Search completed: December 21, 2000, 08:37:54
Job time: 287 sec

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92406903.
RA   Nakai K., Ward A.M., Gannon M., Rifkind A.B.;
RT   "Beta-naphthoflavone induction of a cytochrome P-450 arachidonic acid
RT   epoxygenase in chick embryo liver distinct from the aryl hydrocarbon
RT   hydroxylase and from phenobarbital-induced arachidonate epoxygenase.";
RL   J. Biol. Chem. 267:19503-19512(1992).
SQ   SEQUENCE   22 AA;  2377 MW;  6BA9C990EE33E9C1 CRC64;


  Query Match             22.1%;  Score 24.5;  DB 13;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 2.1e+03;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       5 MDF-GFPKHLLV 15
           ||| | |  ||:
Db       1 MDFLGLPTILLL 12


RESULT  12
O31364
ID   O31364      PRELIMINARY;      PRT;    13 AA.
AC   O31364;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G25;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93700; AAC45534.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1449 MW;  08FAB9BB0E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


RESULT  13
O31365
ID   O31365      PRELIMINARY;      PRT;    13 AA.
AC   O31365;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP90;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93701; AAC45536.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1382 MW;  08EFB4180E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


RESULT  14
O34622
ID   O34622      PRELIMINARY;      PRT;    13 AA.
AC   O34622;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT39 (ILL-1), DN127CL9-2, AND 25015;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DN127CL9-2, AND 25015;
RX   MEDLINE; 95154673.
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U93694; AAC45522.1; -.
DR   EMBL; U04280; AAC45543.1; -.
DR   EMBL; U04282; AAC45541.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1407 MW;  08FAB3930E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


RESULT  15
Q9W5Q8
ID   Q9W5Q8      PRELIMINARY;      PRT;    19 AA.
AC   Q9W5Q8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CG12575 PROTEIN (FRAGMENT).
GN   CG12575.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
```

```
Qy      6 DFGFPKHL 13
          | | || :
Db      3 DIGLPKRV 10


RESULT   7
O94554
ID   O94554      PRELIMINARY;      PRT;    16 AA.
AC   O94554;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CALMODULIN KINASE 2 (FRAGMENT).
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alemany V., Algue R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57982; AAD09466.1; -.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1846 MW;  4A673B1F4C3288D9 CRC64;


  Query Match             22.5%;  Score 25;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FGFP 10
          ||||
Db      5 FGFP 8


RESULT   8
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;


  Query Match             22.5%;  Score 25;  DB 13;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      2 LLQMDFGFPKHLLVD 16
          ||    | |  | :|
Db      1 LLAQGLGXPTALALD 15


RESULT   9
Q9TWH5
ID   Q9TWH5      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   40 KDA GAP JUNCTION PROTEIN (FRAGMENT).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Heliothis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95347000.
RA   Ryerse J.S.;
RT   "Immunocytochemical, electrophoresis, and immunoblot analysis of
RT   Heliothis virescens gap junctions isolated in the presence and absence
RT   of protease inhibitors.";
RL   Cell Tissue Res. 281:179-186(1995).
SQ   SEQUENCE   20 AA;  2304 MW;  A298D3EB3E89586B CRC64;


  Query Match             22.5%;  Score 25;  DB 5;  Length 20;
  Best Local Similarity   29.4%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPKHLLVDF 17
          ::  :| |: : |  ||
Db      3 VIFNIDGGYLEFLTRDF 19


RESULT  10
Q9TTG3
ID   Q9TTG3      PRELIMINARY;      PRT;    20 AA.
AC   Q9TTG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GALACTOCEREBROSIDASE (FRAGMENT).
GN   CALC.
OS   Ateles belzebuth chamek.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Atelinae; Ateles.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Canavez F.C.;
RT   "Gene mapping in Ateles paniscus chamek.";
RL   Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF099175; AAF21849.1; -.
DR   INTERPRO; IPR001286; -.
DR   PFAM; PF02057; Glyco_hydro_59; 1.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2317 MW;  EA3066655966B58F CRC64;


  Query Match             22.5%;  Score 25;  DB 6;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 1.6e+03;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      2 LLQMDFGFPKHLL 14
          | : :||   |:|
Db      6 LFKPNFGASLHIL 18


RESULT  11
Q9PS42
ID   Q9PS42      PRELIMINARY;      PRT;    22 AA.
AC   Q9PS42;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PHENOBARBITAL-INDUCED 48 KDA CYTOCHROME P-450 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
```

```
GN   PETD.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070133; AAD20742.1; -.
FT   NON_TER      21      21
SQ   SEQUENCE   21 AA;  2410 MW;  CC947A212386FE23 CRC64;


  Query Match             25.2%;  Score 28;  DB 2;  Length 21;
  Best Local Similarity   45.5%;  Pred. No. 5.5e+02;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPK 11
           :| ::||  ||
Db       3 ILKKLDFSDPK 13


RESULT   3
Q9UGS1
ID   Q9UGS1      PRELIMINARY;      PRT;    12 AA.
AC   Q9UGS1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ796I17.4 (NOVEL PROTEIN SIMILAR TO GS2) (FRAGMENT).
GN   DJ796I17.4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035398; CAB63074.1; -.
FT   NON_TER      1       1
SQ   SEQUENCE   12 AA;  1430 MW;  AF7740ABECB69AA6 CRC64;


  Query Match             24.3%;  Score 27;  DB 4;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 4.4e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGF 9
           ::|:||  |
Db       1 MILEMDMSF 9


RESULT   4
Q9TRR6
ID   Q9TRR6      PRELIMINARY;      PRT;    19 AA.
AC   Q9TRR6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN, CAP50=CA2+/PHOSPHOLIPID-BINDING PROTEIN
DE   L-14 FRAGMENT.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92250478.
RA   Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT   "A calcyclin-associated protein is a newly identified member of the
RT   Ca2+/phospholipid-binding proteins, annexin family.";
RL   J. Biol. Chem. 267:8919-8924(1992).
SQ   SEQUENCE   19 AA;  2018 MW;  9A54062504B8322E CRC64;


  Query Match             24.3%;  Score 27;  DB 6;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 7.2e+02;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       7 FGFPKHLLVDFLQSLS 22
           ||  :  ::| | | |
Db       2 FGTDEQAIIDXLGSRS 17


RESULT   5
P70861
ID   P70861      PRELIMINARY;      PRT;    21 AA.
AC   P70861;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   THDF (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE; 97312006.
RA   Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT   "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT   sigma 70-like promoter.";
RL   Microbiology 143:1681-1690(1997).
DR   EMBL; U62901; AAB62742.1; -.
FT   NON_TER      21      21
SQ   SEQUENCE   21 AA;  2432 MW;  F33E1EC548BD5B33 CRC64;


  Query Match             23.4%;  Score 26;  DB 2;  Length 21;
  Best Local Similarity   40.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFPKHLL 14
           ::| ||: :|
Db       7 IEFDFPEGIL 16


RESULT   6
O73594
ID   O73594      PRELIMINARY;      PRT;    10 AA.
AC   O73594;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ZAX-2 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WHITE LEGHORN; TISSUE=WHOLE EMBRYOS;
RX   MEDLINE; 98141813.
RA   Peale F.V., Mason K., Hunter A.W., Bothwell M.;
RT   "Multiplex display polymerase chain reaction amplifies and resolves
RT   related sequences sharing a single moderately conserved domain.";
RL   Anal. Biochem. 256:158-168(1998).
DR   EMBL; U34617; AAC36455.1; -.
FT   NON_TER      1       1
FT   NON_TER     10      10
SQ   SEQUENCE   10 AA;  1068 MW;  9A7598276728705A CRC64;


  Query Match             22.5%;  Score 25;  DB 13;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 7.7e+02;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:33:07 ; Search time 37.32 Seconds
(without alignments)
55.044 Million cell updates/sec

Title: US-08-934-367-4
Perfect score: 111
Sequence: 1 LLLQMDFGFPKHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4962

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30.5	27.5	21	2	Q9ZG55	Q9zg55 chlamydia t
2	28	25.2	21	2	Q9X3D0	Q9x3d0 prochloroco
3	27	24.3	12	4	Q9UGS1	Q9ugs1 homo sapien
4	27	24.3	19	6	Q9TRR6	Q9trr6 oryctolagus
5	26	23.4	21	2	P70861	P70861 borrelia bu
6	25	22.5	10	13	O73594	O73594 gallus gall
7	25	22.5	16	3	O94554	O94554 schizosacch
8	25	22.5	19	13	Q9PS70	Q9ps70 gallus gall
9	25	22.5	20	5	Q9TWH5	Q9twh5 heliothis v
10	25	22.5	20	6	Q9TTG3	Q9ttg3 ateles belz
11	24.5	22.1	22	13	Q9PS42	Q9ps42 gallus gall
12	24	21.6	13	2	O31364	O31364 borrelia ga
13	24	21.6	13	2	O31365	O31365 borrelia ga
14	24	21.6	13	2	O34622	O34622 borrelia bu
15	24	21.6	19	5	Q9W5Q8	Q9w5q8 drosophila
16	23.5	21.2	15	2	Q9R4U7	Q9r4u7 acinetobact
17	23	20.7	8	2	O85406	O85406 coxiella bu
18	23	20.7	9	5	O96417	O96417 drosophila
19	23	20.7	10	2	Q9X534	Q9x534 leclercia a
20	23	20.7	10	2	Q9X533	Q9x533 escherichia
21	23	20.7	13	2	O31296	O31296 borrelia af
22	23	20.7	13	2	O34770	O34770 borrelia af
23	23	20.7	13	11	O35758	O35758 rattus norv
24	23	20.7	14	6	Q9TRQ7	Q9trq7 bos taurus
25	23	20.7	17	2	Q9X3I3	Q9x3i3 prochloroco
26	23	20.7	18	13	P82068	P82068 litoria gen
27	23	20.7	20	2	Q9R4F1	Q9r4f1 desulfovibr
28	23	20.7	21	8	Q9ZYB7	Q9zyb7 spinaria sp
29	23	20.7	21	10	Q41496	Q41496 solanum tub
30	23	20.7	21	12	Q80817	Q80817 human t-cel
31	23	20.7	22	1	P71537	P71537 methanosarc
32	23	20.7	22	10	Q9S8E1	Q9s8e1 capsicum an
33	22.5	20.3	17	6	Q9TR11	Q9tr11 bos taurus
34	22	19.8	13	4	Q9UPE7	Q9upe7 homo sapien
35	22	19.8	14	2	Q56945	Q56945 yersinia ps
36	22	19.8	17	11	Q9QUJ4	Q9quj4 mus sp. mep
37	22	19.8	20	6	Q9TRA1	Q9tra1 bos taurus
38	22	19.8	20	12	Q85096	Q85096 punta toro
39	22	19.8	22	6	O02830	O02830 oryctolagus
40	21.5	19.4	22	2	Q52009	Q52009 pseudomonas
41	21	18.9	9	2	Q51765	Q51765 pseudomonas
42	21	18.9	15	11	Q9Z1I4	Q9z1i4 mus musculu
43	21	18.9	17	8	Q9ZYW0	Q9zyw0 betylobraco
44	21	18.9	18	1	Q9UYK7	Q9uyk7 pyrococcus
45	21	18.9	18	6	O97668	O97668 equus cabal

ALIGNMENTS

```
RESULT   1
Q9ZG55
ID   Q9ZG55      PRELIMINARY;      PRT;    21 AA.
AC   Q9ZG55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ATP-BINDING PROTEIN (FRAGMENT).
GN   RECF.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087306; AAD04082.1; -.
KW   ATP-binding.
FT   NON_TER       1      1
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2336 MW;  0185D9AC428276D9 CRC64;


  Query Match            27.5%;  Score 30.5;  DB 2;  Length 21;
  Best Local Similarity  50.0%;  Pred. No. 2.2e+02;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy       2 LLQMDFGFPKHL 13
           :||:   ||||:
Db       8 ILQL-ISFPKHI 18


RESULT   2
Q9X3D0
ID   Q9X3D0      PRELIMINARY;      PRT;    21 AA.
AC   Q9X3D0;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B6/F COMPLEX SUBUNIT IV (FRAGMENT).
```

```
DR   SUBTILIST; BG?????; ????.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  994 MW;  E89C07772042C051 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 10;
  Best Local Similarity   40.0%;  Pred. No. 1.8e+03;
  Matches    2;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 KHLLV 15
           ||:::
Db       2 KHIVI 6


RESULT  13
SODM_CANFA
ID   SODM_CANFA     STANDARD;      PRT;    13 AA.
AC   P54712;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   SUPEROXIDE DISMUTASE [MN], MITOCHONDRIAL (EC 1.15.1.1) (FRAGMENT).
GN   SOD2.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE
CC       CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.
CC   -!- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) = O(2) + H(2)O(2).
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
CC   -!- SIMILARITY: BELONGS TO THE IRON/MANGANESE SUPEROXIDE DISMUTASE
CC       FAMILY.
DR   HSSP; P04179; 1MSD.
DR   HSC-2DPAGE; P54712; DOG.
DR   INTERPRO; IPR001189; -.
DR   PFAM; PF00081; sodfe; 1.
DR   PROSITE; PS00088; SOD_MN; PARTIAL.
KW   Oxidoreductase; Manganese; Mitochondrion.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1476 MW;  9C9651DE8BE0672A CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      11 KHLLVD 16
           || | |
Db       1 KHSLPD 6


RESULT  14
UP21_UPEIN
ID   UP21_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82027;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.1.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST
CC       L.MESENTERIODES, M.LUTEUS AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1927 MW;  328834D77BA353D2 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.6e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKKV 9


RESULT  15
UP25_UPEIN
ID   UP25_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82031;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.5.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST M.LUTEUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1940; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1941 MW;  5E94C6C757B463D9 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.6e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKGV 9


Search completed: December 21, 2000, 08:38:16
Job time: 149 sec
```

```
RA   Johnson L., Norton S., Landau M., Semmes O.J., Sattelberg R.M.,
RA   Jorenby W.H., Hintz M.F.;
RT   "Characterization of a pigment-dispersing hormone in eyestalks of the
RT   fiddler crab Uca pugilator.";
RL   Proc. Natl. Acad. Sci. U.S.A. 82:5319-5322(1985).
RN   [2]
RP   SEQUENCE.
RX   MEDLINE; 93230895.
RA   Loehr J., Klein J., Webster S.G., Dircksen H.;
RT   "Quantification, immunoaffinity purification and sequence analysis of
RT   a pigment-dispersing hormone of the shore crab, Carcinus maenas
RT   (L.).";
RL   Comp. Biochem. Physiol. 104B:699-706(1993).
CC   -!- FUNCTION: CAUSES THE MIGRATION OF THE DISTAL RETINAL PIGMENT
CC       INTO THE PROXIMAL END OF THE PIGMENT CHROMATOPHORE CELLS AND
CC       THUS DECREASES THE AMOUNT OF LIGHT ENTERING THE RETINULAS.
CC   -!- SIMILARITY: TO THE PDH OF OTHER ARTHROPODS.
DR   PIR; A25144; DRUFPD.
KW   Hormone; Amidation.
FT   DOMAIN        6      9       IMPORTANT FOR DRPH ACTIVITY.
FT   MOD_RES      18     18       AMIDATION.
SQ   SEQUENCE   18 AA;  1928 MW;  25D5CE8D016F544E CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 18;
  Best Local Similarity   40.0%;  Pred. No. 2.4e+03;
  Matches    4;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPK 11
           |:    | ||
Db       4 LINSILGLPK 13


RESULT  10
ATPB_PHYPA
ID   ATPB_PHYPA     STANDARD;      PRT;    21 AA.
AC   P80658;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (FRAGMENT).
GN   ATPB.
OS   Physcomitrella patens (Moss).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida;
OC   Funariidae; Funariales; Funariaceae; Physcomitrella.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PROTONEMA;
RX   MEDLINE; 97275459.
RA   Kasten B., Buck F., Nuske J., Reski R.;
RT   "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT   plastid enzymes.";
RL   Planta 201:261-272(1997).
CC   -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC       GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC       SUBUNIT.
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   MENDEL; 13023; PHYpa;atpB;1.
DR   INTERPRO; IPR000194; -.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW   ATP synthesis; Chloroplast; Thylakoid membrane; CF(1);
KW   Hydrolase; ATP-binding; Hydrogen ion transport.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2298 MW;  9558E4F5AC89D81A CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFP 10
           :|  ||
Db      16 LDIXFP 21


RESULT  11
TRYP_PROAT
ID   TRYP_PROAT     STANDARD;      PRT;    21 AA.
AC   P35051;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4) (FRAGMENT).
OS   Protopterus aethiopicus (Marbled lungfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Dipnoi; Lepidosireniformes; Protopteridae; Protopterus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PANCREAS;
RA   Hermodson M.A., Tye R.W., Reeck G.R., Neurath H., Walsh K.A.;
RT   "Comparison of the amino terminal sequences of bovine, dogfish, and
RT   lungfish trypsinogens.";
RL   FEBS Lett. 14:222-224(1971).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   PIR; A27719; A27719.
DR   HSSP; P07288; 1PFA.
DR   INTERPRO; IPR001254; -.
DR   PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.
DR   PROSITE; PS00135; TRYPSIN_SER; PARTIAL.
KW   Hydrolase; Serine protease; Zymogen.
FT   PROPEP        1      7       ACTIVATION PEPTIDE.
FT   CHAIN         8    >21       TRYPSIN.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2454 MW;  88E2FB11D130729D CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 PKH 12
           |||
Db      15 PKH 17


RESULT  12
GSI5_BACSU
ID   GSI5_BACSU     STANDARD;      PRT;    10 AA.
AC   P80861;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   GLUCOSE STARVATION-INDUCIBLE PROTEIN 5 (GSI5) (FRAGMENT).
OS   Bacillus subtilis.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=IS58;
RX   MEDLINE; 97443988.
RA   Antelmann H., Bernhardt J., Schmid R., Mach H., Voelker U.,
RA   Hecker M.;
RT   "First steps from a two-dimensional protein index towards a response-
RT   regulation map for Bacillus subtilis.";
RL   Electrophoresis 18:1451-1463(1997).
CC   -!- INDUCTION: BY GLUCOSE STARVATION.
CC   -!- SIMILARITY: HIGH, TO GLUTAMYL-TRNA REDUCTASE.
```

```
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20      20
SQ   SEQUENCE   20 AA;  2303 MW;  0A33BD34006E5AA6 CRC64;


  Query Match             19.8%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.3e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 DFGFPKH 12
            | | | |
Db      14 DNGMPVH 20


RESULT   6
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;     9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83160932.
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   INTERPRO; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN       1      >9       PENTAXIN.
FT   NON_TER      9       9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             18.9%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 GFPKHLLV 15
            |||   |:
Db       1 GFPGKSLI 8


RESULT   7
TEML_RANTE
ID   TEML_RANTE     STANDARD;      PRT;     11 AA.
AC   P56923;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN L.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES      11      11       AMIDATION.
SQ   SEQUENCE   11 AA;  1194 MW;  1E990549B3372724 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      13 LLVDFLQSL 21
           || : |:||
Db       1 LLPNLLKSL 9


RESULT   8
MDH_SYNY4
ID   MDH_SYNY4      STANDARD;      PRT;     16 AA.
AC   P80460;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Synechocystis sp. (strain PCC 6714).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN   [1]
RP   SEQUENCE.
RA   Naterstad K., Synstad B., Sirevag R.;
RL   Submitted (SEP-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   INTERPRO; IPR001252; -.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      16      16
SQ   SEQUENCE   16 AA;  1780 MW;  61D1896F14E81984 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 2.1e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LQMDFG 8
            |  |||
Db       9 LTTDFG 14


RESULT   9
DRPH_UCAPU
ID   DRPH_UCAPU     STANDARD;      PRT;     18 AA.
AC   P08871;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PIGMENT-DISPERSING HORMONE (PDH) (LIGHT ADAPTING DISTAL RETINAL
DE   PIGMENT HORMONE) (DRPH).
OS   Uca pugilator (Atlantic sand fiddler crab) (Celuca pugilator).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Ocypodoidea; Ocypodidae; Celuca.
RN   [1]
RP   SEQUENCE.
RA   Rao K.R., Riehm J.P., Zahnow C.A., Kleinholz L.H., Tarr G.E.,
```

```
DE   ANGIOTENSINOGEN (FRAGMENT).
GN   AGT.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RA   Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
RT   "The preparation, purification, and amino acid sequence of a
RT   polypeptide renin substrate.";
RL   J. Exp. Med. 106:439-453(1957).
CC   -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC       CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC       CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC       PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC       PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC       BALANCE OF BODY FLUIDS.
CC   -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC   -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR   PIR; A01250; A01250.
DR   INTERPRO; IPR000215; -.
DR   PROSITE; PS00284; SERPIN; PARTIAL.
KW   Vasoconstrictor; Plasma; Serpin.
FT   PEPTIDE       1     10       ANGIOTENSIN I.
FT   PEPTIDE       1      8       ANGIOTENSIN II.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1759 MW;  2E9921F8EEEFBDD7 CRC64;

  Query Match             21.6%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 4.1e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PKHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT   3
DCMM_PSECF
ID   DCMM_PSECF     STANDARD;      PRT;    14 AA.
AC   P19914;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0142; PL0142.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1756 MW;  65583C6D1FB7C25B CRC64;

  Query Match             21.6%;  Score 24;  DB 1;  Length 14;
  Best Local Similarity   23.1%;  Pred. No. 4.1e+02;
  Matches    3;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPKHL 13
           :: : ::  |||:
Db       2 MIPRFEYHAPKHV 14


RESULT   4
PSBP_PINPS
ID   PSBP_PINPS     STANDARD;      PRT;    13 AA.
AC   P81668;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   OXYGEN-EVOLVING ENHANCER PROTEIN 2 (OEE2) (23 KDA SUBUNIT OF OXYGEN
DE   EVOLVING SYSTEM OF PHOTOSYSTEM II) (FRAGMENT).
GN   PSBP.
OS   Pinus pinaster (Maritime pine).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Coniferopsida; Coniferales; Pinaceae; Pinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=NEEDLE;
RX   MEDLINE; 99274088.
RA   Costa P., Pionneau C., Bauw G., Dubos C., Bahrman N., Kremer A.,
RA   Frigerio J.-M., Plomion C.;
RT   "Separation and characterization of needle and xylem maritime pine
RT   proteins.";
RL   Electrophoresis 20:1098-1108(1999).
CC   -!- FUNCTION: ASSOCIATED WITH THE OXYGEN-EVOLVING COMPLEX OF
CC       PHOTOSYSTEM II (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE; ASSOCIATED
CC       WITH THE PHOTOSYSTEM II COMPLEX (BY SIMILARITY).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS PROTEIN
CC       (SPOT N179) IS: 5.9, ITS MW IS: 22 KDA.
CC   -!- SIMILARITY: TO OTHER OEE2 SUBUNITS.
KW   Photosynthesis; Photosystem II; Chloroplast; Thylakoid membrane.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1294 MW;  C6772B0D54D7C44D CRC64;

  Query Match             19.8%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   80.0%;  Pred. No. 8e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 FGFPK 11
           || ||
Db       9 FGAPK 13


RESULT   5
COXN_THUOB
ID   COXN_THUOB     STANDARD;      PRT;    20 AA.
AC   P80980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:47 ; Search time 18.39 Seconds
(without alignments)
38.219 Million cell updates/sec

Title: US-08-934-367-4
Perfect score: 111
Sequence: 1 LLLQMDFGFPKHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1213

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	24	21.6	11	1	CH60_DROME	P35380 drosophila
2	24	21.6	14	1	ANGT_HORSE	P01016 equus cabal
3	24	21.6	14	1	DCMM_PSECF	P19914 pseudomonas
4	22	19.8	13	1	PSBP_PINPS	P81668 pinus pinas
5	22	19.8	20	1	COXN_THUOB	P80980 thunnus obe
6	21	18.9	9	1	SAMP_MUSCA	P19095 mustelus ca
7	20	18.0	11	1	TEML_RANTE	P56923 rana tempor
8	20	18.0	16	1	MDH_SYNY4	P80460 synechocyst
9	20	18.0	18	1	DRPH_UCAPU	P08871 uca pugilat
10	20	18.0	21	1	ATPB_PHYPA	P80658 physcomitre
11	20	18.0	21	1	TRYP_PROAT	P35051 protopterus
12	19	17.1	10	1	GSI5_BACSU	P80861 bacillus su
13	19	17.1	13	1	SODM_CANFA	P54712 canis famil
14	19	17.1	19	1	UP21_UPEIN	P82027 uperoleia i
15	19	17.1	19	1	UP25_UPEIN	P82031 uperoleia i
16	19	17.1	21	1	DCMS_PSECA	P19921 pseudomonas
17	18	16.2	7	1	ALL7_CYDPO	P82158 cydia pomon
18	18	16.2	8	1	ALL5_CALVO	P41841 calliphora
19	18	16.2	8	1	ALL5_CYDPO	P82156 cydia pomon
20	18	16.2	10	1	PNEU_RAT	P21996 rattus norv
21	18	16.2	12	1	GRAR_RANRU	P40754 rana rugosa
22	18	16.2	13	1	ORCK_ORCLI	P37086 orconectes
23	18	16.2	15	1	LMA2_LOCMI	P38497 locusta mig
24	18	16.2	17	1	FLA2_BARBA	P35634 bartonella
25	18	16.2	19	1	HBB2_UROHA	P18992 uromastyx h
26	18	16.2	20	1	CAQS_RAT	P19633 rattus norv
27	18	16.2	21	1	CSPS_STRTR	P81622 streptococc
28	18	16.2	21	1	OMP4_PASHA	P80228 pasteurella
29	18	16.2	21	1	YD90_HAEIN	P45194 haemophilus
30	18	16.2	22	1	FUC1_RAT	P80347 rattus norv
31	18	16.2	22	1	FUC2_RAT	P80348 rattus norv
32	17	15.3	8	1	CCKN_MACEU	P30369 macropus eu
33	17	15.3	9	1	FAR1_CALVO	P41856 calliphora
34	17	15.3	10	1	CAER_LITXA	P56264 litoria xan
35	17	15.3	12	1	OPS3_DROVI	P17645 drosophila
36	17	15.3	12	1	PSP3_PHYPA	P80662 physcomitre
37	17	15.3	13	1	UP71_LITEW	P82050 litoria ewi
38	17	15.3	15	1	DCMM_PSECA	P19920 pseudomonas
39	17	15.3	15	1	LPF_ECOLI	P03057 escherichia
40	17	15.3	16	1	MMPX_SOLTU	P80501 solanum tub
41	17	15.3	17	1	GAST_MACMU	P33714 macaca mula
42	17	15.3	19	1	IRBP_CAVPO	P12666 cavia porce
43	17	15.3	19	1	NUO6_SOLTU	P80729 solanum tub
44	17	15.3	19	1	UP24_UPEIN	P82030 uperoleia i
45	17	15.3	20	1	CPA7_PAPSP	P80055 papio sp. (

ALIGNMENTS

```
RESULT   1
CH60_DROME
ID   CH60_DROME     STANDARD;      PRT;    11 AA.
AC   P35380;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MITOCHONDRIAL MATRIX PROTEIN P1 (60 KDA CHAPERONIN) (HEAT SHOCK
DE   PROTEIN 60) (HSP-60) (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).
GN   MMP-P1 OR HSP60.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VALLECAS; TISSUE=WING IMAGINAL DISK;
RX   MEDLINE; 93272852.
RA   Santaren J.F., van Damme J., Puype M., Vandekerckhove J.,
RA   Garcia-Bellido A.;
RT   "Identification of Drosophila wing imaginal disc proteins by two-
RT   dimensional gel analysis and microsequencing.";
RL   Exp. Cell Res. 206:220-226(1993).
CC   -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
CC       MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC       IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC       REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC       UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR   FLYBASE; FBgn0010375; Mmp-P1.
DR   INTERPRO; IPR001844; -.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW   Chaperone; ATP-binding; Mitochondrion.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1243 MW;  78501A366365A6DB CRC64;


  Query Match              21.6%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity    27.3%;  Pred. No. 3.2e+02;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPK 11
           ::::  :| ||
Db       1 VIIEQSWGSPK 11


RESULT   2
ANGT_HORSE
ID   ANGT_HORSE     STANDARD;      PRT;    14 AA.
AC   P01016;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
```

C;Date: 22-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 15-Jun-1996
C;Accession: PN0175
R;Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.
submitted to JIPID, December 1995
A;Description: Two dimensional electrophoresis of plant proteins and standardization of
A;Reference number: PN0173
A;Accession: PN0175
A;Molecule type: protein
A;Residues: 1-18 <TSU>
A;Experimental source: leaf
C;Keywords: transferase

Query Match 21.6%; Score 24; DB 2; Length 18;
Best Local Similarity 80.0%; Pred. No. 9.8e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 DFGFP 10
           ||| |
Db       5 DFGHP 9
```

RESULT 15
A37984
ADP,ATP carrier protein - yeast (Candida parapsilosis) (fragment)
C;Species: Candida parapsilosis
C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 07-Jun-1996
C;Accession: A37984
R;Guerin, B.; Bukusoglu, C.; Rakotomanana, F.; Wohlrab, H.
J. Biol. Chem. 265, 19736-19741, 1990
A;Title: Mitochondrial phosphate transport. N-ethylmaleimide insensitivity correlates wi
A;Reference number: A37984; MUID:91060585
A;Accession: A37984
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <GUE>

Query Match 21.6%; Score 24; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 1.1e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy      12 HLLVDFL 18
           : |:|||
Db       5 NFLIDFL 11
```

Search completed: December 21, 2000, 08:37:13
Job time: 317 sec

```
  Matches    8;  Conservative    1;  Mismatches    3;  Indels    5;  Gaps    1;

Qy       1 LLLQMDFGFPKHLLVDF 17
           ||| : |     ||| |
Db       3 LLLSLSF-----LLVGF 14


RESULT   9
S03879
6-phosphofructokinase (EC 2.7.1.11) B - rabbit (fragment)
N;Alternate names: phosphofructo-1-kinase B
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 28-Apr-1993
C;Accession: S03879
R;Valaitis, A.P.; Foe, L.G.; Kwiatkowska, D.; Latshaw, S.P.; Kemp, R.G.
Biochim. Biophys. Acta 995, 187-194, 1989
A;Title: The sites of phosphorylation of rabbit brain phosphofructo-1-kinase by cyclic A
A;Reference number: S03878; MUID:89194250
A;Accession: S03879
A;Molecule type: protein
A;Residues: 1-13 <VAL>
A;Note: the sequence from the summary is inconsistent with that from table I and Fig. 4
C;Keywords: glycolysis; phosphotransferase


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   71.4%;  Pred. No. 6.8e+02;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LQMDFGF 9
           | || ||
Db       7 LSMDKGF 13


RESULT  10
A01250
angiotensin precursor - horse (fragment)
C;Species: Equus caballus (domestic horse)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 20-Mar-1998
C;Accession: A92775; A01250
R;Skeggs Jr., L.T.; Kahn, J.R.; Lentz, K.; Shumway, N.P.
J. Exp. Med. 106, 439-453, 1957
A;Reference number: A92775
A;Accession: A92775
A;Molecule type: protein
A;Residues: 1-14 <SKE>
C;Superfamily: antithrombin III
C;Keywords: blood pressure control; hormone; vasoconstrictor
F;1-10/Product: angiotensin I #status experimental <AN1>
F;1-8/Product: angiotensin II #status experimental <AN2>


  Query Match             21.6%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PKHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT  11
PL0142
carbon-monoxide dehydrogenase (EC 1.2.99.2) medium chain - Pseudomonas carboxydoflava (f
C;Species: Pseudomonas carboxydoflava
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Apr-1993
C;Accession: PL0142
R;Kraut, M.; Hugendieck, I.; Herwig, S.; Meyer, O.
Arch. Microbiol. 152, 335-341, 1989
A;Title: Homology and distribution of CO dehydrogenase structural genes in carboxydotrop
A;Reference number: PL0138; MUID:90055678
A;Accession: PL0142
A;Molecule type: protein
A;Residues: 1-14 <KRA>
C;Comment: Carbon-monoxide dehydrogenase consists of three polypeptide chains: large,
C;Keywords: oxidoreductase


  Query Match             21.6%;  Score 24;  DB 2;  Length 14;
  Best Local Similarity   23.1%;  Pred. No. 7.4e+02;
  Matches    3;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPKHL 13
           :: : ::  |||:
Db       2 MIPRFEYHAPKHV 14


RESULT  12
A60834
angiotensin I precursor - dog (fragment)
N;Alternate names: angiotensinogen I
N;Contains: angiotensin I
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Mar-1998
C;Accession: A60834
R;Oliver, J.A.
Hypertension 11, 21-27, 1988
A;Title: Purification and partial characterization of canine angiotensinogen.
A;Reference number: A60834; MUID:88113996
A;Accession: A60834
A;Molecule type: protein
A;Residues: 1-15 <OLI>
C;Superfamily: antithrombin III
C;Keywords: glycoprotein; plasma
F;1-10/Product: angiotensin I #status predicted <MAT>


  Query Match             21.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 8e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PKHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT  13
B31769
T-cell receptor delta-2 chain J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
C;Accession: B31769
R;Loh, E.Y.; Cwirla, S.; Serafini, A.T.; Phillips, J.H.; Lanier, L.L.
Proc. Natl. Acad. Sci. U.S.A. 85, 9714-9718, 1988
A;Title: Human T-cell-receptor delta chain: genomic organization, diversity, and expr
A;Reference number: A94221; MUID:89071766
A;Accession: B31769
A;Molecule type: DNA
A;Residues: 1-17 <LOH>
A;Cross-references: GB:L36386; NID:g540455; PIDN:AAA61108.1; PID:g540456
C;Keywords: T-cell receptor


  Query Match             21.6%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   37.5%;  Pred. No. 9.2e+02;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPKHLLVD 16
           |  |: ||    |:|:
Db       1 LTAQLFFGKGTQLIVE 16


RESULT  14
PN0175
glutathione transferase (EC 2.5.1.18) 1 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
```

hexokinase (EC 2.7.1.1) I peptide III - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C;Accession: A32521
R;Schirch, D.M.; Wilson, J.E.
Arch. Biochem. Biophys. 257, 1-12, 1987
A;Title: Rat brain hexokinase: amino acid sequence at the substrate hexose binding site
A;Reference number: A90080; MUID:87324917
A;Accession: A32521
A;Molecule type: protein
A;Residues: 1-21 <SCH>
C;Superfamily: human hexokinase I; hexokinase homology
C;Keywords: ATP; glycolysis; phosphotransferase

Query Match 24.3%; Score 27; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LQMDFGFPKH 12
| | || |
Db 3 LGFTFSFPXH 12

RESULT 4
S59492
formate dehydrogenase alpha chain - Alcaligenes eutrophus (fragment)
C;Species: Alcaligenes eutrophus
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59492
R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.
Biol. Chem. Hoppe-Seyler 376, 561-568, 1995
A;Title: Structural and immunological studies on the soluble formate dehydrogenase from
A;Reference number: S59492; MUID:96145736
A;Accession: S59492
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRI>

Query Match 23.4%; Score 26; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 3.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QMDFGFP 10
::||| |
Db 6 EIDFGTP 12

RESULT 5
PX0078
alanine dehydrogenase (EC 1.4.1.1) - oscillatoriacean cyanobacterium (fragment)
C;Species: oscillatoriacean cyanobacterium
C;Date: 22-Apr-1995 #sequence_revision 26-May-1995 #text_change 07-May-1999
C;Accession: PX0078
R;Sawa, Y.; Tani, M.; Murata, K.; Shibata, H.; Ochiai, H.
J. Biochem. 116, 995-1000, 1994
A;Title: Purification and characterization of alanine dehydrogenase from a cyanobacteriu
A;Reference number: PX0078; MUID:95204408
A;Accession: PX0078
A;Molecule type: protein
A;Residues: 1-21 <SAW>
A;Note: the source is designated as Phormidium lapideum
C;Comment: This enzyme catalyzes a reversible oxidative deamination of L-alanine to pyru
C;Superfamily: alanine dehydrogenase; alanine dehydrogenase homology
C;Keywords: oxidoreductase

Query Match 23.4%; Score 26; DB 2; Length 21;
Best Local Similarity 44.4%; Pred. No. 5.4e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 MDFGFPKHL 13
|: | || :
Db 1 MEIGVPKEI 9

RESULT 6
T44936
calmodulin kinase 2 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44936
R;Alemany, V.; Algue, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z22873
A;Accession: T44936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-16 <ALE>
A;Cross-references: EMBL:U57982; PIDN:AAD09466.1

Query Match 22.5%; Score 25; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FGFP 10
||||
Db 5 FGFP 8

RESULT 7
A60743
ornithine carbamoyltransferase (EC 2.1.3.3), anabolic - Aeromonas formicans (fragment
C;Species: Aeromonas formicans
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: A60743
R;Tricot, C.; De Coen, J.L.; Momin, P.; Falmagne, P.; Stalon, V.
J. Gen. Microbiol. 135, 2453-2464, 1989
A;Title: Evolutionary relationships among bacterial carbamoyltransferases.
A;Reference number: A60743; MUID:90188287
A;Accession: A60743
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-17 <TRI>
C;Keywords: transferase

Query Match 22.5%; Score 25; DB 2; Length 17;
Best Local Similarity 83.3%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 11 KHLLVD 16
|||| |
Db 1 KHLLKD 6

RESULT 8
B61597
cytochrome P450 AL-2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: B61597
R;Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.
Drug Metab. Dispos. 19, 291-297, 1991
A;Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto
A;Reference number: A61597; MUID:91292910
A;Accession: B61597
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <SHI>

Query Match 22.1%; Score 24.5; DB 2; Length 14;
Best Local Similarity 47.1%; Pred. No. 6.1e+02;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:56 ; Search time 26.84 Seconds
(without alignments)
52.017 Million cell updates/sec

Title: US-08-934-367-4
Perfect score: 111
Sequence: 1 LLLQMDFGFPKHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 4315

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.2	19	2	A39504	octamer-binding pr
2	27	24.3	7	2	S36662	dermorphin (Lys-7)
3	27	24.3	21	2	A32521	hexokinase (EC 2.7
4	26	23.4	15	2	S59492	formate dehydrogen
5	26	23.4	21	2	PX0078	alanine dehydrogen
6	25	22.5	16	2	T44936	calmodulin kinase
7	25	22.5	17	2	A60743	ornithine carbamoy
8	24.5	22.1	14	2	B61597	cytochrome P450 AL
9	24	21.6	13	2	S03879	6-phosphofructokin
10	24	21.6	14	2	A01250	angiotensin precur
11	24	21.6	14	2	PL0142	carbon-monoxide de
12	24	21.6	15	2	A60834	angiotensin I prec
13	24	21.6	17	2	B31769	T-cell receptor de
14	24	21.6	18	2	PN0175	glutathione transf
15	24	21.6	20	2	A37984	ADP,ATP carrier pr
16	23	20.7	14	2	PA0104	protein QF200070 -
17	23	20.7	19	2	I46654	T-cell receptor de
18	23	20.7	20	2	A41439	acid ribonuclease
19	23	20.7	21	2	T07683	proteinase inhibit
20	22	19.8	7	1	A61324	dermorphin - Rohde
21	22	19.8	10	2	S13224	virG protein - Agr
22	22	19.8	11	2	PT0250	Ig heavy chain CRD
23	22	19.8	12	2	C36201	1-aminocyclopropan
24	22	19.8	14	2	A61002	photosystem II oxy
25	22	19.8	15	2	S71306	heat shock protein
26	22	19.8	20	2	S72501	protein kinase C i
27	22	19.8	20	2	S77989	cytochrome-c oxida
28	22	19.8	20	2	S50203	zona pellucida gly
29	21	18.9	9	2	B45796	dihydrolipoamide S
30	21	18.9	9	2	B20569	serum amyloid P-co
31	21	18.9	13	2	S14995	photosystem II oxy
32	21	18.9	14	2	PA0013	photosystem II oxy
33	21	18.9	17	2	C24166	photosystem II ext
34	21	18.9	18	2	H75063	hypothetical prote
35	21	18.9	19	2	B60822	cytochrome P450 UT
36	21	18.9	19	2	D32071	T-cell receptor de
37	21	18.9	20	2	A60822	cytochrome P450 PB
38	21	18.9	20	2	F56046	urinary tract ston
39	21	18.9	20	2	A54077	cytochrome b558 -
40	20	18.0	9	2	I46023	growth hormone rec
41	20	18.0	12	2	S56122	type I DNA methylt
42	20	18.0	13	2	PC1008	40K extracellular
43	20	18.0	14	2	PQ0152	18K iron-sulfur pr
44	20	18.0	15	2	PH1788	T cell receptor al
45	20	18.0	16	2	I57530	gene c-fms protein

ALIGNMENTS

RESULT 1
A39504
octamer-binding protein, Ku-like, 72K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39504
R;May, G.; Sutton, C.; Gould, H.
J. Biol. Chem. 266, 3052-3059, 1991
A;Title: Purification and characterization of Ku-2, an octamer-binding protein relate
A;Reference number: A39504; MUID:91131605
A;Accession: A39504
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <MAY>

Query Match 25.2%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 LQMDFGF 9
           |:|| ||
Db      11 LEMDVGF 17
```

RESULT 2
S36662
dermorphin (Lys-7) [validated] - two-colored leaf frog
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 10-Dec-1993 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000
C;Accession: S36662
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil,
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of
A;Reference number: S21152; MUID:92339502
A;Accession: S36662
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>

Query Match 24.3%; Score 27; DB 2; Length 7;
Best Local Similarity 80.0%; Pred. No. 1.8e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       7 FGFPK 11
           ||:||
Db       3 FGYPK 7
```

RESULT 3
A32521

```
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-133


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 KHLLVDFLQSLS 22
           : || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT  15
US-07-977-696C-36
; Sequence 36, Application US/07977696C
; Patent No. 5792852
;  GENERAL INFORMATION:
;    APPLICANT:  do Couto, Fernando J.R.
;    APPLICANT:  Ceriani Dr., Roberto L.
;    APPLICANT:  Peterson Dr., Jerry A.
;    APPLICANT:  Padlan Dr., Eduardo A.
;    TITLE OF INVENTION:  Analogue Peptides with Specificity
;    TITLE OF INVENTION:  for Carcinomas and Kit and Diagnostic Vaccination
;    TITLE OF INVENTION:  and Therapeutic Methods.
;    NUMBER OF SEQUENCES:  81
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      STREET:  444 South Flower Street, Suite 2000
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS 5.0
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/977,696C
;      FILING DATE:  11-16-92
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amzel Ph.D., Viviana
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38227
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (510) 748-6868
;      TELEFAX:  (510) 748-6688
;      TELEX:  n.a.
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-07-977-696C-36


  Query Match             24.3%;  Score 27;  DB 1;  Length 19;
  Best Local Similarity   41.2%;  Pred. No. 1.6e+02;
  Matches    7;  Conservative    2;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       5 MDFGFPKHLLVDFLQSL 21
           ||||     ||  |: :
Db       1 MDFGLSLVFLVLILKGV 17


Search completed: December 21, 2000, 08:36:43
Job time: 387 sec
```

```
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/940,095
;       FILING DATE:  29-SEP-1997
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  009196-0004-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  133:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  22 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  No. 6004925e
US-08-940-095-133


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 KHLLVDFLQSLS 22
           : || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT  13
US-08-940-093-133
; Sequence 133, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;     APPLICANT:  Dasseux, Jean-Louis
;     APPLICANT:  Sekul, Renate
;     APPLICANT:  Buttner, Klaus
;     APPLICANT:  Cornut, Isabelle
;     APPLICANT:  Metz, Gunther
;     TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;     TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;     NUMBER OF SEQUENCES:  258
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/940,093
;       FILING DATE:  29-SEP-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  009196-0006-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  133:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  22 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  No. 6037323e
US-08-940-093-133


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 KHLLVDFLQSLS 22
           : || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT  14
US-08-940-096-133
; Sequence 133, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;     APPLICANT:  Dasseux, Jean-Louis
;     APPLICANT:  Sekul, Renate
;     APPLICANT:  Buttner, Klaus
;     APPLICANT:  Cornut, Isabelle
;     APPLICANT:  Metz, Gunther
;     TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;     TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;     NUMBER OF SEQUENCES:  258
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Pennie & Edmonds LLP
;       STREET:  1155 Avenue of the Americas
;       CITY:  New York
;       STATE:  NY
;       COUNTRY:  USA
;       ZIP:  10036-2811
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/940,096
;       FILING DATE:  29-SEP-1997
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Coruzzi, Laura A
;       REGISTRATION NUMBER:  30,742
;       REFERENCE/DOCKET NUMBER:  009196-0005-999
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-493-4935
;       TELEFAX:  650-493-5556
;       TELEX:  66141 PENNIE
;   INFORMATION FOR SEQ ID NO:  133:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  22 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
```

```
RESULT  10
US-08-802-981-112
; Sequence 112, Application US/08802981
; Patent No. 6037137
;  GENERAL INFORMATION:
;    APPLICANT:  Komoriya, Akira
;    APPLICANT:  Packard, Beverly S.
;    TITLE OF INVENTION:  Compositions for the Detection of Enzyme
;    TITLE OF INVENTION:  Activity in Biological Samples and Methods of Use Thereof
;    NUMBER OF SEQUENCES:  231
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/802,981
;      FILING DATE:  20-FEB-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  016865-000300US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  112:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /product= "Aib"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /product= "Acp"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  16
;      OTHER INFORMATION:  /product= "Acp"
US-08-802-981-112


  Query Match             25.2%;  Score 28;  DB 3;  Length 21;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPK 11
           |  :|| ||
Db      10 LDAEFGXPK 18


RESULT  11
US-08-802-981-113
; Sequence 113, Application US/08802981
; Patent No. 6037137
;  GENERAL INFORMATION:
;    APPLICANT:  Komoriya, Akira
;    APPLICANT:  Packard, Beverly S.
;    TITLE OF INVENTION:  Compositions for the Detection of Enzyme
;    TITLE OF INVENTION:  Activity in Biological Samples and Methods of Use Thereof
;    NUMBER OF SEQUENCES:  231
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/802,981
;      FILING DATE:  20-FEB-1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hunter, Tom
;      REGISTRATION NUMBER:  38,498
;      REFERENCE/DOCKET NUMBER:  016865-000300US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  113:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  3
;      OTHER INFORMATION:  /product= "Aib"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  4
;      OTHER INFORMATION:  /product= "Acp"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  16
;      OTHER INFORMATION:  /product= "Acp"
US-08-802-981-113


  Query Match             25.2%;  Score 28;  DB 3;  Length 21;
  Best Local Similarity   55.6%;  Pred. No. 1.2e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPK 11
           |  :|| ||
Db      10 LDAEFGXPK 18


RESULT  12
US-08-940-095-133
; Sequence 133, Application US/08940095
; Patent No. 6004925
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    APPLICANT:  Dufourcq, Jean
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,567
;      FILING DATE:
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/061,514
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Grant D.
;      REGISTRATION NUMBER:  31,259
;      REFERENCE/DOCKET NUMBER:  0941.001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-601-2706
;      TELEFAX:  510-655-3542
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-08-370-567-7


  Query Match             25.2%;  Score 28;  DB 1;  Length 17;
  Best Local Similarity   46.2%;  Pred. No. 96;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPKHLLV 15
           | ||    ||:|:
Db       4 LFMDLWHDKHILL 16


RESULT   8
US-08-438-759-7
; Sequence 7, Application US/08438759
; Patent No. 5679782
;  GENERAL INFORMATION:
;    APPLICANT:  Rosenberg, Steven
;    APPLICANT:  Doyle, Michael
;    APPLICANT:  Goodson, Robert
;    TITLE OF INVENTION:  Peptide Inhibitors of Urokinase Receptor
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/438,759
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/370,567
;      FILING DATE:
;      APPLICATION NUMBER:  US/08/061,514
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Grant D.
;      REGISTRATION NUMBER:  31,259
;      REFERENCE/DOCKET NUMBER:  0941.001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-601-2706
;      TELEFAX:  510-655-3542
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
US-08-438-759-7


  Query Match             25.2%;  Score 28;  DB 1;  Length 17;
  Best Local Similarity   46.2%;  Pred. No. 96;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPKHLLV 15
           | ||    ||:|:
Db       4 LFMDLWHDKHILL 16


RESULT   9
PCT-US94-05684-7
; Sequence 7, Application PC/TUS9405684
;  GENERAL INFORMATION:
;    APPLICANT:  Chiron Corporation
;    TITLE OF INVENTION:  Peptide Inhibitors of Urokinase Receptor
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/05684
;      FILING DATE:  19 MAY 1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Green, Grant D.
;      REGISTRATION NUMBER:  31,259
;      REFERENCE/DOCKET NUMBER:  0941.100
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  510-601-2706
;      TELEFAX:  510-655-3542
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    HYPOTHETICAL:  NO
PCT-US94-05684-7


  Query Match             25.2%;  Score 28;  DB 4;  Length 17;
  Best Local Similarity   46.2%;  Pred. No. 96;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPKHLLV 15
           | ||    ||:|:
Db       4 LFMDLWHDKHILL 16
```

```
;      FILING DATE:  18-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/586,536
;      FILING DATE:  21-SEP-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-12/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-827-618A-22


  Query Match             26.1%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 78;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPKHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   5
US-08-483-952A-22
; Sequence 22, Application US/08483952A
; Patent No. 6011139
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/483,952A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,859
;      FILING DATE:  17-SEP-1993
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/716,909
;      FILING DATE:  18-JUN-1991
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/586,536
;      FILING DATE:  21-SEP-1990
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-7/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-483-952A-22


  Query Match             26.1%;  Score 29;  DB 3;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 78;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPKHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   6
US-08-341-018-40
; Sequence 40, Application US/08341018A
; Patent No. 6087323
; GENERAL INFORMATION:
;  APPLICANT: Gywnne, David I.
;  APPLICANT: Mahanthappa, Nagesh K.
;  APPLICANT: Marchionni, Mark A.
;  APPLICANT: Bermingham-McDonogh, Olivia
;  APPLICANT: Goldin, Stanley M.
;  APPLICANT: McBurney, Robert N.
;  TITLE OF INVENTION: USE OF NEUREGULINS AS MODULATORS OF
;  TITLE OF INVENTION: CELLULAR COMMUNICATION
;  FILE REFERENCE: 04585/041001
;  CURRENT APPLICATION NUMBER: US/08/341,018A
;  CURRENT FILING DATE: 1994-11-17
;  NUMBER OF SEQ ID NOS: 87
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 40
;   LENGTH: 9
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-341-018-40


  Query Match             25.2%;  Score 28;  DB 3;  Length 9;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+05;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 KHLLVDFLQ 19
           ||| ::|::
Db       1 KHLGIEFME 9


RESULT   7
US-08-370-567-7
; Sequence 7, Application US/08370567
; Patent No. 5656726
;  GENERAL INFORMATION:
;    APPLICANT:  Rosenberg, Steven
;    APPLICANT:  Doyle, Michael
;    APPLICANT:  Goodson, Robert
;    TITLE OF INVENTION:  Peptide Inhibitors of Urokinase Receptor
;    TITLE OF INVENTION:  Activity
;    NUMBER OF SEQUENCES:  37
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Chiron Corporation
;      STREET:  4560 Horton Street
;      CITY:  Emeryville
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94608
```

```
          ||:| ::  | :|
Db      1 FGYPVYVFGDCVQ 13


RESULT   2
US-08-765-783A-83
; Sequence 83, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-83


  Query Match           26.1%;  Score 29;  DB 2;  Length 9;
  Best Local Similarity  62.5%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 QMDFGFPK 11
          |  ||||:
Db      1 QHHFGFPR 8


RESULT   3
US-08-484-530-22
; Sequence 22, Application US/08484530
; Patent No. 5846740
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisoc
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,530
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-8/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-530-22


  Query Match           26.1%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity  50.0%;  Pred. No. 78;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      5 MDFGFPKHLL 14
            :|| :|  ||
Db      9 IDFHYPNELL 18


RESULT   4
US-08-827-618A-22
; Sequence 22, Application US/08827618A
; Patent No. 5998366
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/827,618A
;      FILING DATE:  09-APR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/485,725
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/716,909
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:16 ; Search time 25.87 Seconds
(without alignments)
14.254 Million cell updates/sec

Title: US-08-934-367-4
Perfect score: 111
Sequence: 1 LLLQMDFGFPKHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 92835

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	27.0	20	1	US-08-199-508-2	Sequence 2, Appli
2	29	26.1	9	2	US-08-765-783A-83	Sequence 83, Appl
3	29	26.1	20	2	US-08-484-530-22	Sequence 22, Appl
4	29	26.1	20	2	US-08-827-618A-22	Sequence 22, Appl
5	29	26.1	20	3	US-08-483-952A-22	Sequence 22, Appl
6	28	25.2	9	3	US-08-341-018-40	Sequence 40, Appl
7	28	25.2	17	1	US-08-370-567-7	Sequence 7, Appli
8	28	25.2	17	1	US-08-438-759-7	Sequence 7, Appli
9	28	25.2	17	4	PCT-US94-05684-7	Sequence 7, Appli
10	28	25.2	21	3	US-08-802-981-112	Sequence 112, App
11	28	25.2	21	3	US-08-802-981-113	Sequence 113, App
12	28	25.2	22	3	US-08-940-095-133	Sequence 133, App
13	28	25.2	22	3	US-08-940-093-133	Sequence 133, App
14	28	25.2	22	3	US-08-940-096-133	Sequence 133, App
15	27	24.3	19	1	US-07-977-696C-36	Sequence 36, Appl
16	27	24.3	19	1	US-08-129-930B-36	Sequence 36, Appl
17	27	24.3	20	2	US-08-733-505A-48	Sequence 48, Appl
18	27	24.3	20	2	US-08-706-741B-83	Sequence 83, Appl
19	27	24.3	20	2	US-08-924-695A-83	Sequence 83, Appl
20	26	23.4	7	1	US-07-923-724-47	Sequence 47, Appl
21	26	23.4	7	2	US-08-609-426A-47	Sequence 47, Appl
22	26	23.4	7	2	US-08-374-652C-37	Sequence 37, Appl
23	26	23.4	11	1	US-08-039-778B-6	Sequence 6, Appli
24	26	23.4	12	3	US-08-329-799-21	Sequence 21, Appl
25	26	23.4	15	1	US-08-208-181A-16	Sequence 16, Appl
26	26	23.4	15	1	US-08-208-181A-22	Sequence 22, Appl
27	26	23.4	15	3	US-08-596-257A-10	Sequence 10, Appl
28	26	23.4	15	3	US-08-860-339-10	Sequence 10, Appl
29	26	23.4	16	1	US-08-346-455B-52	Sequence 52, Appl
30	26	23.4	16	3	US-08-802-981-71	Sequence 71, Appl
31	26	23.4	16	3	US-08-977-221-52	Sequence 52, Appl
32	26	23.4	16	4	PCT-US95-06613-52	Sequence 52, Appl
33	26	23.4	19	1	US-08-238-163-20	Sequence 20, Appl
34	26	23.4	21	3	US-08-802-981-114	Sequence 114, App
35	26	23.4	21	3	US-08-802-981-116	Sequence 116, App
36	26	23.4	22	1	US-08-039-778B-5	Sequence 5, Appli
37	26	23.4	22	5	5248607-2	Patent No. 5248607
38	25.5	23.0	12	1	US-08-467-940-18	Sequence 18, Appl
39	25.5	23.0	12	1	US-08-633-772-18	Sequence 18, Appl
40	25	22.5	6	1	US-08-487-006-65	Sequence 65, Appl
41	25	22.5	6	1	US-08-487-006-76	Sequence 76, Appl
42	25	22.5	6	2	US-08-488-659A-65	Sequence 65, Appl
43	25	22.5	6	2	US-08-488-659A-76	Sequence 76, Appl
44	25	22.5	8	3	US-08-582-776C-29	Sequence 29, Appl
45	25	22.5	8	3	US-08-434-831B-29	Sequence 29, Appl

ALIGNMENTS

```
RESULT 1
US-08-199-508-2
; Sequence 2, Application US/08199508
; Patent No. 5717058
; GENERAL INFORMATION:
;   APPLICANT: Matthews, Maura-Ann H.
;   APPLICANT: Stetler, Gary L.
;   APPLICANT: Anthony-Cahill, Spencer J.
;   APPLICANT: Anderson, David C.
;   TITLE OF INVENTION: Modulators of Gene Expression
;   NUMBER OF SEQUENCES: 57
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Somatogen, Inc.
;     STREET: 5797 Central Avenue
;     CITY: Boulder
;     STATE: Colorado
;     ZIP: 80301
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb storage
;     COMPUTER: Apple Macintosh
;     OPERATING SYSTEM: System 7.0.1
;     SOFTWARE: Microsoft Word 5.0a
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/199,508
;     FILING DATE: February 18, 1994
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/021,536
;     FILING DATE: February 23, 1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: No. 5717058ak, Henry P.
;     REGISTRATION NUMBER: 33200
;     REFERENCE/DOCKET NUMBER: 121 CIP
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 303-541-3322
;     TELEFAX: 303-444-3013
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 20
;     TYPE: amino acid
;     TOPOLOGY: unknown to applicant
;   MOLECULE TYPE: peptide
;   HYPOTHETICAL: yes
US-08-199-508-2


  Query Match 27.0%; Score 30; DB 1; Length 20;
  Best Local Similarity 38.5%; Pred. No. 53;
  Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 7 FGFPKHLLVDFLQ 19
```

```
CC   effective against antibiotic-resistant strains, and may replace
CC   antibiotic therapy or increase its effectiveness. The present sequence
CC   is shown in the specification.
XX
SQ   Sequence   22 AA;


  Query Match             27.9%;  Score 31;  DB 19;  Length 22;
  Best Local Similarity   54.5%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      7 FGFPKHLLVDF 17
          : || |  |||
Db      4 yqfpghnsvdf 14


RESULT  14
R61276
ID   R61276 standard; peptide; 20 AA.
XX
AC   R61276;
XX
DT   26-APR-1995  (first entry)
XX
DE   Transactivating protein, Tax, peptide fragment #2.
XX
KW   HTLV; transactivating protein; Tax; regulator; host cell; transcription;
KW   protein-protein interaction; inhibition; virus; regulation;
KW   binding site; target molecule.
XX
OS   Human T-cell leukemia Virus.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   20
FT                   /note= "Amidated C-terminal"
XX
PN   WO9419473-A.
XX
PD   01-SEP-1994.
XX
PF   23-FEB-1994;   94WO-IB00020.
XX
PR   23-FEB-1993;   93US-0021536.
XX
PA   (SOMA-) SOMATOGEN INC.
XX
PI   Anderson DC,  Anthony-Cahill SJ,  Matthews MH,  Stetler GL;
XX
DR   WPI; 1994-294336/36.
XX
PT   Regulators of cellular gene transcription - useful in pharmaceutical
PT   compsns. for preventing or treating diseases due to unregulated DNA
PT   DNA transcription
XX
PS   Disclosure; Page 29; 132pp; English.
XX
CC   The sequences given in R6127-316 represent peptides which are derived
CC   from the HTLV transactivating protein, Tax.  These peptides were
CC   preferably synthesised as C-terminal amides, and are not acylated on
CC   the N-terminus.  These peptides may be used as regulators of host
CC   cell transcription, as they modulate other protein-protein inter-
CC   actions which regulate transcription in a cell.  These peptides
CC   inhibit gene transcription in a viral host cell which is subject to
CC   regulation by proteins or factors that originate from a virus.
CC   These peptides specifically act by blocking a binding site, or by
CC   mimicing or simulating binding characteristics of the target molecule.
XX
SQ   Sequence   20 AA;


  Query Match             27.0%;  Score 30;  DB 15;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      7 FGFPKHLLVDFLQ 19
          ||:| ::  | :|
Db      1 fgypvyvfgdcvq 13


RESULT  15
W41190
ID   W41190 standard; peptide; 20 AA.
XX
AC   W41190;
XX
DT   13-MAY-1998  (first entry)
XX
DE   Tax protein fragment.
XX
KW   Tax protein fragment; HTLV-1; human T-cell leukaemia virus; hepatitis B;
KW   human lymphotropic virus; trans-activating factor inhibitor; therapy;
KW   Tax-dependent transcription; viral infection; genetic disorder;
KW   homozygous familial hypercholesterolaemia; cancer.
XX
OS   Human lymphotropic virus.
XX
PN   US5717058-A.
XX
PD   10-FEB-1998.
XX
PF   18-FEB-1994;   94US-0199508.
XX
PR   18-FEB-1994;   94US-0199508.
PR   23-FEB-1993;   93US-0021536.
XX
PA   (SOMA-) SOMATOGEN INC.
XX
PI   Anderson DC,  Anthony-Cahill SJ,  Matthews MH,  Stetler GL;
XX
DR   WPI; 1998-144847/13.
XX
PT   Trans-activating factor inhibitory peptide(s) - comprising fragments
PT   of Tax protein
XX
PS   Example 1; column 41-42; 43pp; English.
XX
CC   This sequence represents a fragment of the Tax protein of human T-cell
CC   leukaemia virus (also known as human lymphotropic virus). This sequence
CC   is used in the trans-activating factor inhibitor of the invention, that
CC   inhibits Tax-dependent transcription. The inhibitors can be used to
CC   inhibit transcription of DNA that codes for proteins that can be harmful
CC   to mammals, especially humans, e.g. to treat viral infections (e.g.
CC   HTLV-I infections or hepatitis B), genetic disorders (e.g. homozygous
CC   familial hypercholesterolaemia) or cancer.
XX
SQ   Sequence   20 AA;


  Query Match             27.0%;  Score 30;  DB 19;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 1.4e+02;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      7 FGFPKHLLVDFLQ 19
          ||:| ::  | :|
Db      1 fgypvyvfgdcvq 13


Search completed: December 21, 2000, 08:36:12
Job time: 471 sec
```

```
PT   Phosphorylated receptor peptide binds STAT transcription factor -
PT   has a Tyr-contg. region of an intracellular domain of, e.g. a
PT   cytokine, for screening for agents which bind to the receptor
XX
PS   Claim 4; Page 39; 55pp; English.
XX
CC   This sequence represents the fragment Lys433 to Asp448 of the
CC   intracellular region of the interferon-gamma receptor protein.  This
CC   peptide represents a Tyr-containing region which is capable of being
CC   phosphorylated, and can then bind a cytoplasmic transcription factor
CC   of the signal transducers and activators of transcription (STAT)
CC   family.  This peptide specifically binds to STAT1alpha.  Receptor
CC   peptides such as this, may be used to screen for molecules which
CC   interact with it, and for molecules which inhibit or promote an
CC   interaction between the receptor peptide and the identified molecule.
XX
SQ   Sequence   16 AA;


  Query Match             28.4%;  Score 31.5;  DB 17;  Length 16;
  Best Local Similarity   63.6%;  Pred. No. 62;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    1;  Gaps    1;

Qy      7 FGFPK-HLLVD 16
          ||: | |:|||
Db      6 fgydkphvlvd 16


RESULT  12
W65666
ID   W65666 standard; peptide; 21 AA.
XX
AC   W65666;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Fibronectin binding protein-derived peptide #10.
XX
KW   microbial surface components recognising adhesive matrix molecule;
KW   MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
KW   antibiotic; bacterial infection; antibiotic-resistant strain.
XX
OS   Synthetic.
OS   Staphylococcus aureus.
XX
PN   WO9831389-A2.
XX
PD   23-JUL-1998.
XX
PF   21-JAN-1998;   98WO-US01222.
XX
PR   21-JAN-1997;   97US-0036139.
XX
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Hoeoek M,  House-pompeo KL,  Joh D,  Mcgavin MJ,  Patti JM;
PI   Speziale P;
XX
DR   WPI; 1998-413816/35.
XX
PT   Antibody that binds to fibronectin-binding protein, preventing its
PT   binding to fibronectin - used to treat or prevent bacterial
PT   infection, especially by Staphylococci and Streptococci
XX
PS   Example 2; Page 92; 201pp; English.
XX
CC   The invention relates to antibodies that bind to a fibronectin-binding
CC   domain of a fibronectin-binding protein, and inhibit binding of the
CC   protein to fibronectin. Also claimed are: (1) isolated peptides of a
CC   fibronectin-binding protein that do not bind to fibronectin; (2) fusion
CC   protein containing at least one peptide of a fibronectin-binding protein
CC   linked to a second amino acid sequence; (3) nucleic acid encoding
CC   the isolated peptide of (1). Antibodies, the isolated peptides of
CC   (1) and the nucleic acids are all useful for immunisation (active or
CC   passive) and (by inhibiting binding of bacteria to fibronectin) for
CC   preventing or treating infection in humans or other animals, particularly
CC   by staphylococci or streptococci, e.g. meningitis, otitis media,
CC   pneumonia, endocarditis, mastitis in cattle, abortion in horses and many
CC   others. Since the antibodies block binding of bacteria, they should be
CC   effective against antibiotic-resistant strains, and may replace
CC   antibiotic therapy or increase its effectiveness. Sequences W65659-68
CC   represent synthetic epitopes from the fibronectin binding domains DU and
CC   D1-D4 of the S. aureus fnbA gene.
XX
SQ   Sequence   21 AA;


  Query Match             27.9%;  Score 31;  DB 19;  Length 21;
  Best Local Similarity   54.5%;  Pred. No. 1e+02;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      7 FGFPKHLLVDF 17
          : || |  |||
Db      4 yqfpghnsvdf 14


RESULT  13
W65669
ID   W65669 standard; peptide; 22 AA.
XX
AC   W65669;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Peptide #13.
XX
KW   microbial surface components recognising adhesive matrix molecule;
KW   MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
KW   antibiotic; bacterial infection; antibiotic-resistant strain.
XX
OS   Synthetic.
XX
PN   WO9831389-A2.
XX
PD   23-JUL-1998.
XX
PF   21-JAN-1998;   98WO-US01222.
XX
PR   21-JAN-1997;   97US-0036139.
XX
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Hoeoek M,  House-pompeo KL,  Joh D,  Mcgavin MJ,  Patti JM;
PI   Speziale P;
XX
DR   WPI; 1998-413816/35.
XX
PT   Antibody that binds to fibronectin-binding protein, preventing its
PT   binding to fibronectin - used to treat or prevent bacterial
PT   infection, especially by Staphylococci and Streptococci
XX
PS   Example 3; Page 93; 201pp; English.
XX
CC   The invention relates to antibodies that bind to a fibronectin-binding
CC   domain of a fibronectin-binding protein, and inhibit binding of the
CC   protein to fibronectin. Also claimed are: (1) isolated peptides of a
CC   fibronectin-binding protein that do not bind to fibronectin; (2) fusion
CC   protein containing at least one peptide of a fibronectin-binding protein
CC   linked to a second amino acid sequence; (3) nucleic acid encoding
CC   the isolated peptide of (1). Antibodies, the isolated peptides of
CC   (1) and the nucleic acids are all useful for immunisation (active or
CC   passive) and (by inhibiting binding of bacteria to fibronectin) for
CC   preventing or treating infection in humans or other animals, particularly
CC   by staphylococci or streptococci, e.g. meningitis, otitis media,
CC   pneumonia, endocarditis, mastitis in cattle, abortion in horses and many
CC   others. Since the antibodies block binding of bacteria, they should be
```

```
PR   17-NOV-1997;   97US-0066090.
PR   17-NOV-1997;   97US-0066094.
PR   17-NOV-1997;   97US-0066095.
PR   17-NOV-1997;   97US-0066089.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Carter KC,  Ebner R,  Endress GA,  Feng P,  Janat F;
PI   Kyaw H,  Lafleur DW,  Moore PA,  Ni J,  Olsen HS,  Rosen CA;
PI   Ruben SM,  Shi Y,  Soppet DR,  Wei Y;
XX
DR   WPI; 1999-337740/28.
DR   N-PSDB; X85016.
XX
PT   New human secreted proteins and coding sequences useful for treating
PT   disorders of the immune system and hyperproliferative disorders
XX
PS   Disclosure; Page 119; 507pp; English.
XX
CC   This sequence represents a secreted human protein encoded by the gene
CC   clone detailed in the descriptor line.  The gene can be used to generate
CC   fusion proteins by linking to the gene to a human immunoglobulin Fc
CC   portion (e.g. X84924) for increasing the stability of the fused protein
CC   as compared to the human protein only.
CC   The invention relates to 125 novel genes and their fragments (nucleic
CC   acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
CC   are useful for preventing, treating or ameliorating  medical conditions
CC   e.g. by protein or gene therapy.  Also, pathological conditions can be
CC   diagnosed by determining the amount of the new polypeptides in a sample
CC   or by determining the presence of mutations in the new polynucleotides.
CC   Specific uses are described for each of the 125 polynucleotides, based on
CC   which tissues they are most highly expressed in (see X84933 for described
CC   uses).
XX
SQ   Sequence   14 AA;


  Query Match             28.8%;  Score 32;  DB 20;  Length 14;
  Best Local Similarity   61.5%;  Pred. No. 44;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      10 PKHLLVDFLQSLS 22
           |    | ||||||
Db       1 phssrvsflqsls 13


RESULT  10
W65690
ID   W65690 standard; peptide; 21 AA.
XX
AC   W65690;
XX
DT   16-OCT-1998  (first entry)
XX
DE   Fibronectin binding protein-derived peptide #34.
XX
KW   microbial surface components recognising adhesive matrix molecule;
KW   MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
KW   antibiotic; bacterial infection; antibiotic-resistant strain.
XX
OS   Synthetic.
OS   Staphylococcus aureus.
XX
PN   WO9831389-A2.
XX
PD   23-JUL-1998.
XX
PF   21-JAN-1998;   98WO-US01222.
XX
PR   21-JAN-1997;   97US-0036139.
XX
PA   (TEXA ) UNIV TEXAS A & M SYSTEM.
XX
PI   Hoeoek M,  House-pompeo KL,  Joh D,  Mcgavin MJ,  Patti JM;
PI   Speziale P;
XX
DR   WPI; 1998-413816/35.
XX
PT   Antibody that binds to fibronectin-binding protein, preventing its
PT   binding to fibronectin - used to treat or prevent bacterial
PT   infection, especially by Staphylococci and Streptococci
XX
PS   Example 8; Page 101; 201pp; English.
XX
CC   The invention relates to antibodies that bind to a fibronectin-binding
CC   domain of a fibronectin-binding protein, and inhibit binding of the
CC   protein to fibronectin. Also claimed are: (1) isolated peptides of a
CC   fibronectin-binding protein that do not bind to fibronectin; (2) fusion
CC   protein containing at least one peptide of a fibronectin-binding protein
CC   linked to a second amino acid sequence; (3) nucleic acid encoding
CC   the isolated peptide of (1). Antibodies, the isolated peptides of
CC   (1) and the nucleic acids are all useful for immunisation (active or
CC   passive) and (by inhibiting binding of bacteria to fibronectin) for
CC   preventing or treating infection in humans or other animals, particularly
CC   by staphylococci or streptococci, e.g. meningitis, otitis media,
CC   pneumonia, endocarditis, mastitis in cattle, abortion in horses and many
CC   others. Since the antibodies block binding of bacteria, they should be
CC   effective against antibiotic-resistant strains, and may replace
CC   antibiotic therapy or increase its effectiveness. Sequences W65670-90
CC   represent a series of synthetic peptides based on the D3 repeat of
CC   S. aureus fibronectin binding protein A. They were synthesised to contain
CC   a proline residue at each position throught the sequence (ie a proline
CC   scan).
XX
SQ   Sequence   21 AA;


  Query Match             28.8%;  Score 32;  DB 19;  Length 21;
  Best Local Similarity   54.5%;  Pred. No. 70;
  Matches    6;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FGFPKHLLVDF 17
           : || |  |||
Db       4 yqfpphnsvdf 14


RESULT  11
R98414
ID   R98414 standard; peptide; 16 AA.
XX
AC   R98414;
XX
DT   18-FEB-1997  (first entry)
XX
DE   Interferon-gamma receptor peptide, Lys433-Asp448.
XX
KW   Interferon-gamma; IFN; receptor protein; phosphorylated;
KW   cytoplasmic transcription factor; STAT1alpha; STAT2;
KW   signal transducers and activators of transcription; STAT.
XX
OS   Homo sapiens.
XX
PN   WO9620211-A1.
XX
PD   04-JUL-1996.
XX
PF   22-DEC-1995;   95WO-US16988.
XX
PR   23-DEC-1994;   94AU-0000249.
XX
PA   (LUDW-) LUDWIG INST CANCER RES.
XX
PI   Harpur AG,  Lackmann M,  Oates AC,  Wilks AF;
XX
DR   WPI; 1996-321801/32.
XX
```

```
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; N30158.
XX
PT   Multi:class hybrid interferon poly:peptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 17; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it.  Pref. the
CC   AA sequence consists of alpha and beta interferons.  Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
CC   N30160, P30227).  Alternativeley IF1 is the amino terminal end of a
CC   beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC   1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC   resp.) (see N30156, P30223).  In the examples plasmids pGW5 and
CC   pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,
CC   N30157).  HinfI was used to digest the DNA sequences in the region
CC   of significant handicaps (see N30153, N30154, N30158, N30159), and
CC   the restriction fragments were ligated to form hybrid DNA.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             28.8%;  Score 32;  DB 4;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 37;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 DFGFPK 11
          |||||:
Db      2 dfgfpq 7
```

```
RESULT   8
P80053
ID   P80053 standard; protein; 12 AA.
XX
AC   P80053;
XX
DT   17-NOV-1990  (first entry)
XX
DE   Sequence of human interferon (huIFN) alpha-61A gene around AA 40.
XX
KW   Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW   antitumour; antiviral; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note="Residue 40"
XX
PN   US4758428-A.
XX
PD   19-JUL-1988.
XX
PF   15-JUL-1985;   85US-0755265.
XX
PR   15-JUL-1985;   85US-0755265.
PR   19-JAN-1983;   83CA-0419758.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1988-219882/31.
DR   N-PSDB; n80050.
XX
PT   Multi-class hybrid interferon polypeptide(s) -
PT   having sequence from interferon-alpha-1 and sequence from
PT   interferon-beta-1 for restricted activity
XX
PS   Example; Fig 17; 24pp; English.
XX
CC   Multi-class hybrid IFN polypeptides having an AA sequence composed
CC   of 2 distinct subsequences are claimed. The plasmids used in the
CC   construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC   pDM101/trp/beta-1. Assembly of the palpha61A plasmid invovled replacing
CC   the DNA fragment encoding the 23 AA signal polypeptide of preinterferon
CC   with a 120BP EcoRI/Sau3A promoter fragment E.coli trp promoter, operator,
CC   and trp leader ribosome binding site preceoperator, encoding an ATG
CC   initiation codon and using HindIII site that was inserted, 59 nucleotides
CC   3'-end of the TGA translational stop codon, to insert the gene into the
CC   plasmid pBW11 (a deriv of pBR322 having a deletion between the HindIII
CC   and PvuII sites). The complete DNA sequence of the promoter and gene
CC   fragments inserted between the EcoRI and HindIII sites of pBW11 is shown
CC   in n80049. The hybrid gene was constructed by taking advantage of the
CC   homologies between huIFN alpha-61A & huIFN beta-1 at around AA 40 of both
CC   proteins. The DNA sequence 5'-proximal to the DdeI restriction enzyme
CC   cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC   3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC   two genes while preserving the translational reading frame of both.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             28.8%;  Score 32;  DB 9;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 37;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 DFGFPK 11
          |||||:
Db      2 dfgfpq 7
```

```
RESULT   9
Y27814
ID   Y27814 standard; Protein; 14 AA.
XX
AC   Y27814;
XX
DT   30-JUL-1999  (first entry)
XX
DE   Human secreted protein encoded by gene No. 84.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9924836-A1.
XX
PD   20-MAY-1999.
XX
PF   04-NOV-1998;   98WO-US23435.
XX
PR   17-NOV-1997;   97US-0066100.
PR   07-NOV-1997;   97US-0064900.
PR   07-NOV-1997;   97US-0064908.
PR   07-NOV-1997;   97US-0064911.
PR   07-NOV-1997;   97US-0064912.
PR   07-NOV-1997;   97US-0064983.
PR   07-NOV-1997;   97US-0064984.
PR   07-NOV-1997;   97US-0064985.
PR   07-NOV-1997;   97US-0064987.
PR   07-NOV-1997;   97US-0064988.
```

```
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
XX
SQ   Sequence   16 AA;


  Query Match             71.2%;  Score 79;  DB 21;  Length 16;
  Best Local Similarity   93.8%;  Pred. No. 1.5e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFPKHLLVDFLQSLS 22
           ||||:|||||||||||
Db       1 fgfpehllvdflqsls 16


RESULT   6
W24294
ID   W24294 standard; peptide; 11 AA.
XX
AC   W24294;
XX
DT   17-OCT-1997  (first entry)
XX
DE   Human/Rabbit CETP common peptide.
XX
KW   Immune response; high density lipoprotein; HDL; cholesterol; human;
KW   serum; epitope; cholesteryl ester transfer protein; CETP; rabbit.
XX
OS   Oryctolagus cuniculus.
OS   Homo sapiens.
XX
PN   WO9639168-A1.
XX
PD   12-DEC-1996.
XX
PF   05-JUN-1996;   96WO-US09143.
XX
PR   06-JUN-1995;   95US-0482454.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW,  Carlo DJ,  Kwoh DY;
XX
DR   WPI; 1997-042849/04.
XX
PT   Stimulating an immune response to increase high density lipoprotein
PT   - avoids repeated administration of toxic drugs to lower cholesteryl
PT   ester transfer protein levels
XX
PS   Claim 5; Page 16; 26pp; English.
XX
CC   The sequences given in W24292-94 were used in the method of the
CC   invention to stimulate an immune response to increase high density
CC   lipoprotein (HDL) cholesterol in a mammal exhibiting low levels of
CC   serum HDL.  These peptides represent immunogenic epitopes of
CC   cholesteryl ester transfer protein (CETP).  The method utilises
CC   the body's own immune system to lower CETP levels, thereby increasing
CC   the level of beneficial HDL cholesterol, preferably in serum.  The
CC   method avoids the problems associated with the repeated administration
CC   of drugs which have undesirable side effects.  This peptide represents
CC   a region of CETP which is common to both human and rabbit proteins.
XX
SQ   Sequence   11 AA;


  Query Match             47.7%;  Score 53;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.015;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 HLLVDFLQSLS 22
           ||||||||||||
Db       1 hllvdflqsls 11


RESULT   7
P30225
ID   P30225 standard; Protein; 12 AA.
XX
AC   P30225;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -alpha-61A around amino acid 40.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /label= AA No. 40
XX
PN   WO8302461-A.
XX
PD   21-JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU-) CETUS CORP.
XX
```

Y91230
ID Y91230 standard; peptide; 16 AA.
XX
AC Y91230;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:108.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 74.8%; Score 83; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 3.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FGFPKHLLVDFLQSLS 22
||||||||||||||||
Db 1 fgfpkhllvdflqsls 16

RESULT 5
Y91229
ID Y91229 standard; peptide; 16 AA.
XX
AC Y91229;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:107.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 50; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport

```
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             100.0%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.7e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPKHLLVDFLQSLS 22
          ||||||||||||||||||||||
Db      1 lllqmdfgfpkhllvdflqsls 22


RESULT   2
Y13815
ID   Y13815 standard; peptide; 22 AA.
XX
AC   Y13815;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 75; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             96.4%;  Score 107;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 7.3e-11;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPKHLLVDFLQSLS 22
          ||||||||||:||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   3
Y13821
ID   Y13821 standard; peptide; 22 AA.
XX
AC   Y13821;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Human CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Homo sapiens.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 88; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the human
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             96.4%;  Score 107;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 7.3e-11;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPKHLLVDFLQSLS 22
          ||||||||||:||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   4
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:28:21 ; Search time 34.96 Seconds
(without alignments)
21.518 Million cell updates/sec

Title: US-08-934-367-4
Perfect score: 111
Sequence: 1 LLLQMDFGFPKHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 134459

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	111	100.0	22	20	Y13809	Rabbit CETP immuno
2	107	96.4	22	20	Y13815	Rabbit CETP immuno
3	107	96.4	22	20	Y13821	Human CETP immunog
4	83	74.8	16	21	Y91230	Human cholesteryl
5	79	71.2	16	21	Y91229	Human cholesteryl
6	53	47.7	11	18	W24294	Human/Rabbit CETP
7	32	28.8	12	4	P30225	Sequence of interf
8	32	28.8	12	9	P80053	Sequence of human
9	32	28.8	14	20	Y27814	Human secreted pro
10	32	28.8	21	19	W65690	Fibronectin bindin
11	31.5	28.4	16	17	R98414	Interferon-gamma r
12	31	27.9	21	19	W65666	Fibronectin bindin
13	31	27.9	22	19	W65669	Peptide #13. Synt
14	30	27.0	20	15	R61276	Transactivating pr
15	30	27.0	20	19	W41190	Tax protein fragme
16	29	26.1	20	16	R72269	Glutamic acid deca
17	29	26.1	20	21	Y59570	GAD65 fragment, pe
18	29	26.1	22	17	R95899	Fragment #4 of 7-a
19	28	25.2	9	15	R46908	GGF segment D'. H
20	28	25.2	9	15	R55648	GGF segment D'. H
21	28	25.2	9	16	R67236	Bovine glial cell
22	28	25.2	9	17	W09354	Bovine glial growt
23	28	25.2	9	17	R96068	Human glial growth
24	28	25.2	9	17	R86700	Bovine glial growt
25	28	25.2	9	19	W72564	Dengue virus type-
26	28	25.2	9	20	Y26568	Bovine neuregulin
27	28	25.2	15	20	W73425	Human secreted pro
28	28	25.2	17	16	R67556	Antagonist (7) of
29	28	25.2	20	19	W82211	Fluorogenic protea
30	28	25.2	21	18	W38080	PPPPY motif contai
31	28	25.2	21	19	W82186	Fluorogenic protea
32	28	25.2	21	19	W82187	Fluorogenic protea
33	28	25.2	22	20	Y19066	Lecithin:cholester
34	28	25.2	22	20	Y18812	Lecithin:cholester
35	28	25.2	22	20	Y18549	Lecithin:cholester
36	28	25.2	22	20	Y19320	Lecithin:cholester
37	27	24.3	9	17	R88341	Adenomatous polypo
38	27	24.3	11	18	W10733	Activated matrix m
39	27	24.3	12	16	R78515	Synthetic HTLV pep
40	27	24.3	12	18	W28864	HTLV-1 derived pep
41	27	24.3	12	20	Y32796	Mammalian prolacti
42	27	24.3	13	19	W76044	Hydra head activat
43	27	24.3	13	19	W62713	Streptococcus pneu
44	27	24.3	14	13	R21617	Sequence encoded b
45	27	24.3	19	16	R70458	VH sequence of ant

ALIGNMENTS

RESULT 1
Y13809
ID Y13809 standard; peptide; 22 AA.
XX
AC Y13809;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Example 1; Page 73; 99pp; English.
XX

CC uraemia, diabetes, Cushing's syndrome, cachexias, eating disorders, AIDS,
CC after stress and during neuromuscular disease.
XX
SQ Sequence 12 AA;

Query Match 27.9%; Score 31; DB 20; Length 12;
Best Local Similarity 60.0%; Pred. No. 70;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

```
Qy      11 EHLLVDFLQS 20
           | |||:||::
Db       1 edllveflen 10
```

Search completed: December 21, 2000, 08:36:13
Job time: 472 sec

```
CC   apolipoprotein A-I (apoA-I) agonist (B) that is a peptide, or analog,
CC   which forms an amphipathic alpha-helix in presence of lipids. (A),
CC   optionally as a complex with lipids, and host cells that contain (A),
CC   are useful for gene therapy, or prevention, of diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) and apoA-I
CC   deficiency, hypertriglyceridemia and metabolic syndrome, also to treat
CC   endotoxemia (septic shock). Host cells containing (A) can also be used
CC   to study the role of apoA-I in lipid metabolism. (B) can be used
CC   diagnostically, e.g. to measure serum HDL (particularly its
CC   subpopulation involved in retrograde cholesterol transport) and for
CC   imaging the circulatory system or HDL accumulations at fatty streaks.
CC   The present sequence represents a peptide from the present invention.
XX
SQ   Sequence   22 AA;

  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  14
Y19320
ID   Y19320 standard; Peptide; 22 AA.
XX
AC   Y19320;
XX
DT   14-JUL-1999  (first entry)
XX
DE   Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW   Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW   human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW   cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW   high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW   septic shock.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916459-A1.
XX
PD   08-APR-1999.
XX
PF   28-SEP-1998;   98WO-US20327.
XX
PR   29-SEP-1997;   97US-0940095.
XX
PA   (BUTT/) BUTTNER K.
PA   (CORN/) CORNUT I.
PA   (DASS/) DASSEUX J.
PA   (DUFO/) DUFOURCQ J.
PA   (METZ/) METZ G.
PA   (SEKU/) SEKUL R.
XX
PI   Buttner K,  Cornut I,  Dasseux J,  Dufourcq J,  Metz G;
PI   Sekul R;
XX
DR   WPI; 1999-277035/23.
XX
PT   Peptide agonists of apolipoprotein A-I
XX
PS   Claim 19; Page 157; 280pp; English.
XX
CC   The present invention describes an agonist (A) of apolipoprotein A-I
CC   (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC   amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC   complexes, are used to treat or prevent diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC   deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC   treating septic shock. When labeled, (A) can also be used diagnostically
CC   to measure serum levels of HDL, in particular the HDL subpopulation that
CC   is involved in retrograde cholesterol transport, also to image HDL at
CC   e.g. atherosclerotic streaks, and to raise antibodies. Y19188 to Y19441
CC   represent lecithin:cholesterol acyltransferase (LCAT) activity
CC   exhibiting core peptides, which are apoA-I agonists.
XX
SQ   Sequence   22 AA;

  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  15
Y32796
ID   Y32796 standard; peptide; 12 AA.
XX
AC   Y32796;
XX
DT   09-NOV-1999  (first entry)
XX
DE   Mammalian prolactin receptor derived polypeptide.
XX
KW   Signal transduction; proteolytic cleavage; prolactin receptor;
KW   proteosome binding site; muscle wasting; renal tubular defect; uraemia;
KW   diabetes; Cushing's disease; eating disorder; AIDS; ubiquitin;
KW   growth hormone deficiency; inhibitor.
XX
OS   Oryctolagus cuniculus.
OS   Rattus sp.
OS   Mus sp.
XX
PN   EP943624-A1.
XX
PD   22-SEP-1999.
XX
PF   12-MAR-1998;   98EP-0200799.
XX
PR   12-MAR-1998;   98EP-0200799.
XX
PA   (UYUT-) RIJKSUNIV UTRECHT.
XX
DR   WPI; 1999-510568/43.
XX
PT   Controlling the availability and/or signal transduction capability
PT   of a cell surface receptor, useful for treating growth hormone
PT   deficiencies
XX
PS   Disclosure; Page 9; 36pp; English.
XX
CC   This sequence is a polypeptide from the mammalian prolactin receptor.
CC   Sequences (Y32794-Y32823) are examples of polypeptide sequences found at
CC   or near the ubiquitin/proteosome complex binding site located on the
CC   intracellular part of a cell surface receptor. These sequences are used
CC   in a method for controlling the availability and signal transduction
CC   capability of a cell surface receptor by administering an inhibitor that
CC   is capable of inhibiting proteolytic cleavage of the receptor. Inhibition
CC   of this proteolytic cleavage results in the receptors being present on
CC   the surface for longer and therefore signalling for longer to the
CC   interior of the cell. This increases the sensitivity of cells to any
CC   hormones which might be present. The inhibitor is either derived from,
CC   competes with or binds to a polypeptide sequence of which sequences
CC   Y32794-Y32823 are examples. The inhibitor may be used to treat muscle
CC   wasting, associated with disorders such as renal tubular defects,
```

```
PA   (METZ/) METZ G.
PA   (SEKU/) SEKUL R.
XX
PI   Buttner K,  Cornut I,  Dasseux J,  Metz G,  Sekul R;
XX
DR   WPI; 1999-277034/23.
XX
PT   Peptide agonists of apolipoprotein A-I
XX
PS   Example; Page 114; 254pp; English.
XX
CC   The present invention describes an agonist (A) of apolipoprotein A-I
CC   (apoA-I) which is a 15-29 residue peptide, or analog, that forms an
CC   amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC   complexes, are used to treat or prevent diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC   deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC   treating septic shock. When labeled, (A) can also be used diagnostically
CC   to measure serum levels of HDL, in particular the HDL subpopulation that
CC   is involved in retrograde cholesterol transport, also to image HDL at
CC   e.g. atherosclerotic streaks, and to raise antibodies. Y18934 to Y19187
CC   represent lecithin:cholesterol acyltransferase (LCAT) activity
CC   exhibiting core peptides, which are apoA-I agonists.
XX
SQ   Sequence   22 AA;


  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  12
Y18812
ID   Y18812 standard; Peptide; 22 AA.
XX
AC   Y18812;
XX
DT   09-JUL-1999  (first entry)
XX
DE   Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW   Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW   human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW   cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW   high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW   septic shock.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916408-A2.
XX
PD   08-APR-1999.
XX
PF   28-SEP-1998;   98WO-US20328.
XX
PR   29-SEP-1997;   97US-0940093.
XX
PA   (BUTT/) BUTTNER K.
PA   (CORN/) CORNUT I.
PA   (DASS/) DASSEUX J.
PA   (METZ/) METZ G.
PA   (SEKU/) SEKUL R.
XX
PI   Buttner K,  Cornut I,  Dasseux J,  Metz G,  Sekul R;
XX
DR   WPI; 1999-277031/23.
XX
PT   Peptide agonists of apolipoprotein A-I
XX
PS   Example; Page 111; 152pp; English.
XX
CC   The present invention describes an agonist (A) of apolipoprotein A-I
CC   (apoA-I) which is a 14-22 residue peptide, or analog, that forms an
CC   amphipathic alpha-helix in presence of lipids. (A), and their lipid
CC   complexes, are used to treat or prevent diseases associated with
CC   dyslipidemia, specifically hypercholesterolaemia, cardiovascular disease,
CC   atherosclerosis, restenosis, HDL (high density lipoprotein) or apoA-I
CC   deficiency; hypertriglyceridemia and metabolic syndrome, also for
CC   treating septic shock. When labeled, (A) can also be used diagnostically
CC   to measure serum levels of HDL, in particular the HDL subpopulation that
CC   is involved in retrograde cholesterol transport, also to image HDL at
CC   e.g. atherosclerotic streaks, and to raise antibodies. Y18680 to Y18933
CC   represent lecithin:cholesterol acyltransferase (LCAT) activity
CC   exhibiting core peptides, which are apoA-I agonists.
XX
SQ   Sequence   22 AA;


  Query Match             28.8%;  Score 32;  DB 20;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 94;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 erlledllqaln 18


RESULT  13
Y18549
ID   Y18549 standard; Peptide; 22 AA.
XX
AC   Y18549;
XX
DT   09-JUL-1999  (first entry)
XX
DE   Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW   Gene therapy; apolipoprotein A-I; agonist; dyslipidemic disorder;
KW   ApoA-I; cardiovascular disease; atherosclerosis; restenosis; LCAT;
KW   hyperlipidemia; septic shock; lecithin:cholesterol acyltransferase.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9916409-A2.
XX
PD   08-APR-1999.
XX
PF   28-SEP-1998;   98WO-US20329.
XX
PR   29-SEP-1997;   97US-0940136.
XX
PA   (BUTT/) BUTTNER K.
PA   (CORN/) CORNUT I.
PA   (DASS/) DASSEUX J.
PA   (DUFO/) DUFOURCQ J.
PA   (METZ/) METZ G.
PA   (SEKU/) SEKUL R.
XX
PI   Buttner K,  Cornut I,  Dasseux J,  Dufourcq J,  Metz G;
PI   Sekul R;
XX
DR   WPI; 1999-254921/21.
XX
PT   Nucleic acid encoding apolipoprotein A-I agonist peptides
XX
PS   Claim 18; Page 166; 232pp; English.
XX
CC   The present invention describes a nucleic acid (A) encoding an
```

PR 17-NOV-1997; 97US-0066090.
PR 17-NOV-1997; 97US-0066094.
PR 17-NOV-1997; 97US-0066095.
PR 17-NOV-1997; 97US-0066089.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;
PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y;
XX
DR WPI; 1999-337740/28.
DR N-PSDB; X85016.
XX
PT New human secreted proteins and coding sequences useful for treating
PT disorders of the immune system and hyperproliferative disorders
XX
PS Disclosure; Page 119; 507pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. X84924) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 125 novel genes and their fragments (nucleic
CC acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 125 polynucleotides, based on
CC which tissues they are most highly expressed in (see X84933 for described
CC uses).
XX
SQ Sequence 14 AA;

Query Match 29.7%; Score 33; DB 20; Length 14;
Best Local Similarity 61.5%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 10 PEHLLVDFLQSLS 22
| | |||||
Db 1 phssrvsflqsls 13

RESULT 10
W65690
ID W65690 standard; peptide; 21 AA.
XX
AC W65690;
XX
DT 16-OCT-1998 (first entry)
XX
DE Fibronectin binding protein-derived peptide #34.
XX
KW microbial surface components recognising adhesive matrix molecule;
KW MSCRAMM; fibronectin; fibronectin-binding protein; antibody; epitope;
KW antibiotic; bacterial infection; antibiotic-resistant strain.
XX
OS Synthetic.
OS Staphylococcus aureus.
XX
PN WO9831389-A2.
XX
PD 23-JUL-1998.
XX
PF 21-JAN-1998; 98WO-US01222.
XX
PR 21-JAN-1997; 97US-0036139.
XX
PA (TEXA) UNIV TEXAS A & M SYSTEM.
XX
PI Hoeoek M, House-pompeo KL, Joh D, Mcgavin MJ, Patti JM;
PI Speziale P;
XX
DR WPI; 1998-413816/35.
XX
PT Antibody that binds to fibronectin-binding protein, preventing its
PT binding to fibronectin - used to treat or prevent bacterial
PT infection, especially by Staphylococci and Streptococci
XX
PS Example 8; Page 101; 201pp; English.
XX
CC The invention relates to antibodies that bind to a fibronectin-binding
CC domain of a fibronectin-binding protein, and inhibit binding of the
CC protein to fibronectin. Also claimed are: (1) isolated peptides of a
CC fibronectin-binding protein that do not bind to fibronectin; (2) fusion
CC protein containing at least one peptide of a fibronectin-binding protein
CC linked to a second amino acid sequence; (3) nucleic acid encoding
CC the isolated peptide of (1). Antibodies, the isolated peptides of
CC (1) and the nucleic acids are all useful for immunisation (active or
CC passive) and (by inhibiting binding of bacteria to fibronectin) for
CC preventing or treating infection in humans or other animals, particularly
CC by staphylococci or streptococci, e.g. meningitis, otitis media,
CC pneumonia, endocarditis, mastitis in cattle, abortion in horses and many
CC others. Since the antibodies block binding of bacteria, they should be
CC effective against antibiotic-resistant strains, and may replace
CC antibiotic therapy or increase its effectiveness. Sequences W65670-90
CC represent a series of synthetic peptides based on the D3 repeat of
CC S. aureus fibronectin binding protein A. They were synthesised to contain
CC a proline residue at each position throught the sequence (ie a proline
CC scan).
XX
SQ Sequence 21 AA;

Query Match 28.8%; Score 32; DB 19; Length 21;
Best Local Similarity 54.5%; Pred. No. 90;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 7 FGFPEHLLVDF 17
: || | |||
Db 4 yqfpphnsvdf 14

RESULT 11
Y19066
ID Y19066 standard; Peptide; 22 AA.
XX
AC Y19066;
XX
DT 09-JUL-1999 (first entry)
XX
DE Lecithin:cholesterol acyltransferase activation exhibiting peptide #133.
XX
KW Apolipoprotein A-I; agonist; dyslipidemic disorder; dyslipidemia;
KW human; lecithin:cholesterol acyltransferase; LCAT; hypercholesterolaemia;
KW cardiovascular disease; atherosclerosis; restenosis; HDL; apoA-I;
KW high density lipoprotein; hypertriglyceridemia; metabolic syndrome;
KW septic shock.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9916458-A1.
XX
PD 08-APR-1999.
XX
PF 28-SEP-1998; 98WO-US20326.
XX
PR 29-SEP-1997; 97US-0940096.
XX
PA (BUTT/) BUTTNER K.
PA (CORN/) CORNUT I.
PA (DASS/) DASSEUX J.

```
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; N30158.
XX
PT   Multi:class hybrid interferon poly:peptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 17; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it.  Pref. the
CC   AA sequence consists of alpha and beta interferons.  Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
CC   N30160, P30227).  Alternativeley IF1 is the amino terminal end of a
CC   beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC   1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC   resp.) (see N30156, P30223).  In the examples plasmids pGW5 and
CC   pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,
CC   N30157).  HinfI was used to digest the DNA sequences in the region
CC   of significant handicaps (see N30153, N30154, N30158, N30159), and
CC   the restriction fragments were ligated to form hybrid DNA.
XX
SQ   Sequence   12 AA;


  Query Match             29.7%;  Score 33;  DB 4;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 34;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 DFGFPE 11
           |||||:
Db       2 dfgfpq 7


RESULT   8
P80053
ID   P80053 standard; protein; 12 AA.
XX
AC   P80053;
XX
DT   17-NOV-1990  (first entry)
XX
DE   Sequence of human interferon (huIFN) alpha-61A gene around AA 40.
XX
KW   Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW   antitumour; antiviral; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note="Residue 40"
XX
PN   US4758428-A.
XX
PD   19-JUL-1988.
XX
PF   15-JUL-1985;   85US-0755265.
XX
PR   15-JUL-1985;   85US-0755265.
PR   19-JAN-1983;   83CA-0419758.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1988-219882/31.
DR   N-PSDB; n80050.
XX
PT   Multi-class hybrid interferon polypeptide(s) -
PT   having sequence from interferon-alpha-1 and sequence from
PT   interferon-beta-1 for restricted activity
XX
PS   Example; Fig 17; 24pp; English.
XX
CC   Multi-class hybrid IFN polypeptides having an AA sequence composed
CC   of 2 distinct subsequences are claimed. The plasmids used in the
CC   construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC   pDM101/trp/beta-1. Assembly of the palpha61A plasmid invovled replacing
CC   the DNA fragment encoding the 23 AA signal polypeptide of preinterferon
CC   with a 120BP EcoRI/Sau3A promoter fragment E.coli trp promoter, operator,
CC   and trp leader ribosome binding site preceoperator, encoding an ATG
CC   initiation codon and using HindIII site that was inserted, 59 nucleotides
CC   3'-end of the TGA translational stop codon, to insert the gene into the
CC   plasmid pBW11 (a deriv of pBR322 having a deletion between the HindIII
CC   and PvuII sites). The complete DNA sequence of the promoter and gene
CC   fragments inserted between the EcoRI and HindIII sites of pBW11 is shown
CC   in n80049. The hybrid gene was constructed by taking advantage of the
CC   homologies between huIFN alpha-61A & huIFN beta-1 at around AA 40 of both
CC   proteins. The DNA sequence 5'-proximal to the DdeI restriction enzyme
CC   cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC   3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC   two genes while preserving the translational reading frame of both.
XX
SQ   Sequence   12 AA;


  Query Match             29.7%;  Score 33;  DB 9;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 34;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 DFGFPE 11
           |||||:
Db       2 dfgfpq 7


RESULT   9
Y27814
ID   Y27814 standard; Protein; 14 AA.
XX
AC   Y27814;
XX
DT   30-JUL-1999  (first entry)
XX
DE   Human secreted protein encoded by gene No. 84.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
PN   WO9924836-A1.
XX
PD   20-MAY-1999.
XX
PF   04-NOV-1998;   98WO-US23435.
XX
PR   17-NOV-1997;   97US-0066100.
PR   07-NOV-1997;   97US-0064900.
PR   07-NOV-1997;   97US-0064908.
PR   07-NOV-1997;   97US-0064911.
PR   07-NOV-1997;   97US-0064912.
PR   07-NOV-1997;   97US-0064983.
PR   07-NOV-1997;   97US-0064984.
PR   07-NOV-1997;   97US-0064985.
PR   07-NOV-1997;   97US-0064987.
PR   07-NOV-1997;   97US-0064988.
```

```
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
XX
SQ   Sequence   16 AA;


  Query Match             71.2%;  Score 79;  DB 21;  Length 16;
  Best Local Similarity   93.8%;  Pred. No. 3.5e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFLQSLS 22
           ||||:|||||||||||
Db       1 fgfpkhllvdflqsls 16


RESULT   6
W24294
ID   W24294 standard; peptide; 11 AA.
XX
AC   W24294;
XX
DT   17-OCT-1997  (first entry)
XX
DE   Human/Rabbit CETP common peptide.
XX
KW   Immune response; high density lipoprotein; HDL; cholesterol; human;
KW   serum; epitope; cholesteryl ester transfer protein; CETP; rabbit.
XX
OS   Oryctolagus cuniculus.
OS   Homo sapiens.
XX
PN   WO9639168-A1.
XX
PD   12-DEC-1996.
XX
PF   05-JUN-1996;   96WO-US09143.
XX
PR   06-JUN-1995;   95US-0482454.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW,  Carlo DJ,  Kwoh DY;
XX
DR   WPI; 1997-042849/04.
XX
PT   Stimulating an immune response to increase high density lipoprotein
PT   - avoids repeated administration of toxic drugs to lower cholesteryl
PT   ester transfer protein levels
XX
PS   Claim 5; Page 16; 26pp; English.
XX
CC   The sequences given in W24292-94 were used in the method of the
CC   invention to stimulate an immune response to increase high density
CC   lipoprotein (HDL) cholesterol in a mammal exhibiting low levels of
CC   serum HDL.  These peptides represent immunogenic epitopes of
CC   cholesteryl ester transfer protein (CETP).  The method utilises
CC   the body's own immune system to lower CETP levels, thereby increasing
CC   the level of beneficial HDL cholesterol, preferably in serum.  The
CC   method avoids the problems associated with the repeated administration
CC   of drugs which have undesirable side effects.  This peptide represents
CC   a region of CETP which is common to both human and rabbit proteins.
XX
SQ   Sequence   11 AA;


  Query Match             47.7%;  Score 53;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.025;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 HLLVDFLQSLS 22
           ||||||||||||
Db       1 hllvdflqsls 11


RESULT   7
P30225
ID   P30225 standard; Protein; 12 AA.
XX
AC   P30225;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -alpha-61A around amino acid 40.
XX
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /label= AA No. 40
XX
PN   WO8302461-A.
XX
PD   21-JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU-) CETUS CORP.
XX
```

Y91229
ID Y91229 standard; peptide; 16 AA.
XX
AC Y91229;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:107.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 50; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 74.8%; Score 83; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 8.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FGFPEHLLVDFLQSLS 22
||||||||||||||||
Db 1 fgfpehllvdflqsls 16

RESULT 5
Y91230
ID Y91230 standard; peptide; 16 AA.
XX
AC Y91230;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:108.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport

```
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             100.0%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||||||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   2
Y13821
ID   Y13821 standard; peptide; 22 AA.
XX
AC   Y13821;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Human CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Homo sapiens.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 88; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the human
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             100.0%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 5.6e-11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||||||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   3
Y13809
ID   Y13809 standard; peptide; 22 AA.
XX
AC   Y13809;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Example 1; Page 73; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             96.4%;  Score 107;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 2.3e-10;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHLLVDFLQSLS 22
          ||||||||||:||||||||||
Db      1 lllqmdfgfpkhllvdflqsls 22


RESULT   4
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:13 ; Search time 34.96 Seconds
(without alignments)
21.518 Million cell updates/sec

Title: US-08-934-367-34
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 134459

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	111	100.0	22	20	Y13815	Rabbit CETP immuno
2	111	100.0	22	20	Y13821	Human CETP immunog
3	107	96.4	22	20	Y13809	Rabbit CETP immuno
4	83	74.8	16	21	Y91229	Human cholesteryl
5	79	71.2	16	21	Y91230	Human cholesteryl
6	53	47.7	11	18	W24294	Human/Rabbit CETP
7	33	29.7	12	4	P30225	Sequence of interf
8	33	29.7	12	9	P80053	Sequence of human
9	33	29.7	14	20	Y27814	Human secreted pro
10	32	28.8	21	19	W65690	Fibronectin bindin
11	32	28.8	22	20	Y19066	Lecithin:cholester
12	32	28.8	22	20	Y18812	Lecithin:cholester
13	32	28.8	22	20	Y18549	Lecithin:cholester
14	32	28.8	22	20	Y19320	Lecithin:cholester
15	31	27.9	12	20	Y32796	Mammalian prolacti
16	31	27.9	21	19	W65666	Fibronectin bindin
17	31	27.9	22	19	W65669	Peptide #13. Synt
18	30	27.0	20	15	R61276	Transactivating pr
19	30	27.0	20	19	W41190	Tax protein fragme
20	30	27.0	22	17	R95899	Fragment #4 of 7-a
21	29	26.1	11	20	W99442	Interleukin-2 rece
22	29	26.1	11	21	Y82915	Peptide exhibiting
23	29	26.1	13	18	W10883	MAb anti-HBsAg bin
24	29	26.1	19	16	W13199	Fragment of p53 bi
25	29	26.1	19	19	W57353	Human WBP1 immunog
26	29	26.1	20	16	R72269	Glutamic acid deca
27	29	26.1	20	21	Y59570	GAD65 fragment, pe
28	29	26.1	21	21	Y69793	Human interleukin-
29	28	25.2	9	19	W72564	Dengue virus type-
30	28	25.2	10	15	Y38102	Hepatitis B virus-
31	28	25.2	10	20	Y45670	Immunogenic peptid
32	28	25.2	13	19	W62713	Streptococcus pneu
33	28	25.2	14	13	R21617	Sequence encoded b
34	28	25.2	15	19	W45817	Peptide recognised
35	28	25.2	15	19	W45612	Peptide recognised
36	28	25.2	21	18	W38080	PPPPY motif contai
37	28	25.2	22	20	Y19067	Lecithin:cholester
38	28	25.2	22	20	Y18813	Lecithin:cholester
39	28	25.2	22	20	Y18550	Lecithin:cholester
40	28	25.2	22	20	Y19321	Lecithin:cholester
41	28	25.2	22	20	W96827	Nucleic acid bindi
42	27.5	24.8	16	17	R98414	Interferon-gamma r
43	27	24.3	9	17	R88341	Adenomatous polypo
44	27	24.3	12	16	R78515	Synthetic HTLV pep
45	27	24.3	12	18	W28864	HTLV-1 derived pep

ALIGNMENTS

RESULT 1
Y13815
ID Y13815 standard; peptide; 22 AA.
XX
AC Y13815;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Disclosure; Page 75; 99pp; English.
XX

KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WO9924836-A1.
XX
PD 20-MAY-1999.
XX
PF 04-NOV-1998; 98WO-US23435.
XX
PR 17-NOV-1997; 97US-0066100.
PR 07-NOV-1997; 97US-0064900.
PR 07-NOV-1997; 97US-0064908.
PR 07-NOV-1997; 97US-0064911.
PR 07-NOV-1997; 97US-0064912.
PR 07-NOV-1997; 97US-0064983.
PR 07-NOV-1997; 97US-0064984.
PR 07-NOV-1997; 97US-0064985.
PR 07-NOV-1997; 97US-0064987.
PR 07-NOV-1997; 97US-0064988.
PR 17-NOV-1997; 97US-0066090.
PR 17-NOV-1997; 97US-0066094.
PR 17-NOV-1997; 97US-0066095.
PR 17-NOV-1997; 97US-0066089.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Carter KC, Ebner R, Endress GA, Feng P, Janat F;
PI Kyaw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Wei Y;
XX
DR WPI; 1999-337740/28.
DR N-PSDB; X85016.
XX
PT New human secreted proteins and coding sequences useful for treating
PT disorders of the immune system and hyperproliferative disorders
XX
PS Disclosure; Page 119; 507pp; English.
XX
CC This sequence represents a secreted human protein encoded by the gene
CC clone detailed in the descriptor line. The gene can be used to generate
CC fusion proteins by linking to the gene to a human immunoglobulin Fc
CC portion (e.g. X84924) for increasing the stability of the fused protein
CC as compared to the human protein only.
CC The invention relates to 125 novel genes and their fragments (nucleic
CC acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
CC are useful for preventing, treating or ameliorating medical conditions
CC e.g. by protein or gene therapy. Also, pathological conditions can be
CC diagnosed by determining the amount of the new polypeptides in a sample
CC or by determining the presence of mutations in the new polynucleotides.
CC Specific uses are described for each of the 125 polynucleotides, based on
CC which tissues they are most highly expressed in (see X84933 for described
CC uses).
XX
SQ Sequence 14 AA;

Query Match 24.6%; Score 33; DB 20; Length 14;
Best Local Similarity 61.5%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 14 PEHLLVDFLQSLS 26
| | ||||||
Db 1 phssrvsflqsls 13

Search completed: December 21, 2000, 08:38:46
Job time: 150 sec

```
KW   Hybrid interferon; antiviral; therapy; cancer; tumour.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /label= AA No. 40
XX
PN   WO8302461-A.
XX
PD   21-JUL-1983.
XX
PF   18-JAN-1983;   83WO-0900607.
XX
PR   19-JAN-1982;   82US-0340782.
PR   03-FEB-1983;   83US-0463574.
PR   15-JUL-1985;   85US-0755265.
XX
PA   (CETU-) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1983-723186/30.
DR   N-PSDB; N30158.
XX
PT   Multi:class hybrid interferon poly:peptide(s) - with restricted
PT   antiviral and cell growth regulatory activities
XX
PS   Example; Fig 17; 61pp; English.
XX
CC   The inventors claim a multiclass hybrid interferon polypeptide and a
CC   DNA unit having a nucleotide sequence which encodes it.  Pref. the
CC   AA sequence consists of alpha and beta interferons.  Pref. IF1 is
CC   (i) the 1-73 AA seq. of HuIFN-alpha-1 (and IF2 is the 74-166 AA seq.
CC   of HuIFN-beta-1) (see N30155, P30222); or (ii) the 1-41 AA seq. of
CC   HuIFN-alpha-61A (and IF2 is the 43-166 AA seq. of HuIFN-beta-1) (see
CC   N30160, P30227).  Alternativeley IF1 is the amino terminal end of a
CC   beta-IF and IF2 is the carboxy terminal of an alpha-IF (esp. the
CC   1-73 seq. of HuIFN-beta-1 and the 74-167 seq. of HuIFN-alpha-1
CC   resp.) (see N30156, P30223).  In the examples plasmids pGW5 and
CC   pDM101/trp/beta-1 and p-alpha-61A were used (see N30151, N30152,
CC   N30157).  HinfI was used to digest the DNA sequences in the region
CC   of significant handicaps (see N30153, N30154, N30158, N30159), and
CC   the restriction fragments were ligated to form hybrid DNA.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             24.6%;  Score 33;  DB 4;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 DFGFPE 15
           |||||:
Db       2 dfgfpq 7
```

```
RESULT  14
P80053
ID   P80053 standard; protein; 12 AA.
XX
AC   P80053;
XX
DT   17-NOV-1990  (first entry)
XX
DE   Sequence of human interferon (huIFN) alpha-61A gene around AA 40.
XX
KW   Alpha-beta hybrid interferon; multi-class hybrid interferon;
KW   antitumour; antiviral; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 7
FT                   /note="Residue 40"
XX
PN   US4758428-A.
XX
PD   19-JUL-1988.
XX
PF   15-JUL-1985;   85US-0755265.
XX
PR   15-JUL-1985;   85US-0755265.
PR   19-JAN-1983;   83CA-0419758.
XX
PA   (CETU ) CETUS CORP.
XX
PI   Mark DF,  Creasey AA;
XX
DR   WPI; 1988-219882/31.
DR   N-PSDB; n80050.
XX
PT   Multi-class hybrid interferon polypeptide(s) -
PT   having sequence from interferon-alpha-1 and sequence from
PT   interferon-beta-1 for restricted activity
XX
PS   Example; Fig 17; 24pp; English.
XX
CC   Multi-class hybrid IFN polypeptides having an AA sequence composed
CC   of 2 distinct subsequences are claimed. The plasmids used in the
CC   construction of huIFN-alpha-61A-beta-1 hybrid are plasmids palpha61A and
CC   pDM101/trp/beta-1. Assembly of the palpha61A plasmid invovled replacing
CC   the DNA fragment encoding the 23 AA signal polypeptide of preinterferon
CC   with a 120BP EcoRI/Sau3A promoter fragment E.coli trp promoter, operator,
CC   and trp leader ribosome binding site preceoperator, encoding an ATG
CC   initiation codon and using HindIII site that was inserted, 59 nucleotides
CC   3'-end of the TGA translational stop codon, to insert the gene into the
CC   plasmid pBW11 (a deriv of pBR322 having a deletion between the HindIII
CC   and PvuII sites). The complete DNA sequence of the promoter and gene
CC   fragments inserted between the EcoRI and HindIII sites of pBW11 is shown
CC   in n80049. The hybrid gene was constructed by taking advantage of the
CC   homologies between huIFN alpha-61A & huIFN beta-1 at around AA 40 of both
CC   proteins. The DNA sequence 5'-proximal to the DdeI restriction enzyme
CC   cutting site of the huIFN alpha-61A DNA is ligated to the DNA sequence
CC   3'-proximal to the site of huIFN beta-1, to create a fusion of the
CC   two genes while preserving the translational reading frame of both.
XX
SQ   Sequence   12 AA;
```

```
  Query Match             24.6%;  Score 33;  DB 9;  Length 12;
  Best Local Similarity   83.3%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      10 DFGFPE 15
           |||||:
Db       2 dfgfpq 7
```

```
RESULT  15
Y27814
ID   Y27814 standard; Protein; 14 AA.
XX
AC   Y27814;
XX
DT   30-JUL-1999  (first entry)
XX
DE   Human secreted protein encoded by gene No. 84.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
```

```
Qy      11 FGFPEHLLVDFLQSLS 26
           ||||:||||||||||
Db       1 fgfpkhllvdflqsls 16


RESULT  11
W24294
ID   W24294 standard; peptide; 11 AA.
XX
AC   W24294;
XX
DT   17-OCT-1997  (first entry)
XX
DE   Human/Rabbit CETP common peptide.
XX
KW   Immune response; high density lipoprotein; HDL; cholesterol; human;
KW   serum; epitope; cholesteryl ester transfer protein; CETP; rabbit.
XX
OS   Oryctolagus cuniculus.
OS   Homo sapiens.
XX
PN   WO9639168-A1.
XX
PD   12-DEC-1996.
XX
PF   05-JUN-1996;   96WO-US09143.
XX
PR   06-JUN-1995;   95US-0482454.
XX
PA   (IMMU-) IMMUNE RESPONSE CORP.
XX
PI   Brostoff SW,  Carlo DJ,  Kwoh DY;
XX
DR   WPI; 1997-042849/04.
XX
PT   Stimulating an immune response to increase high density lipoprotein
PT   - avoids repeated administration of toxic drugs to lower cholesteryl
PT   ester transfer protein levels
XX
PS   Claim 5; Page 16; 26pp; English.
XX
CC   The sequences given in W24292-94 were used in the method of the
CC   invention to stimulate an immune response to increase high density
CC   lipoprotein (HDL) cholesterol in a mammal exhibiting low levels of
CC   serum HDL.  These peptides represent immunogenic epitopes of
CC   cholesteryl ester transfer protein (CETP).  The method utilises
CC   the body's own immune system to lower CETP levels, thereby increasing
CC   the level of beneficial HDL cholesterol, preferably in serum.  The
CC   method avoids the problems associated with the repeated administration
CC   of drugs which have undesirable side effects.  This peptide represents
CC   a region of CETP which is common to both human and rabbit proteins.
XX
SQ   Sequence   11 AA;


  Query Match             39.6%;  Score 53;  DB 18;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 0.035;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 HLLVDFLQSLS 26
           |||||||||||
Db       1 hllvdflqsls 11


RESULT  12
Y12067
ID   Y12067 standard; Protein; 25 AA.
XX
AC   Y12067;
XX
DT   18-JUN-1999  (first entry)
XX
XX
DE   Human 5' EST secreted protein SEQ ID NO: 380.
XX
KW   Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW   forensic; gene therapy; chromosome mapping; signal peptide;
KW   upstream regulatory sequence; cytokine activity; cell proliferation;
KW   differentiation; haematopoiesis regulation; tissue growth regulation;
KW   reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW   thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS   Homo sapiens.
XX
PN   WO9906554-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01238.
XX
PR   01-AUG-1997;   97US-0905134.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153784/13.
DR   N-PSDB; X40900.
XX
PT   New nucleic acids encoding human secreted proteins - obtained from
PT   cDNA libraries prepared from kidney, fetal kidney, dystrophic
PT   muscle, muscle and heart tissue
XX
PS   Claim 34; Page 500-501; 622pp; English.
XX
CC   X40826 to X41093 represent 5' expressed sequence tags (ESTs) for human
CC   secreted proteins, and encode the proteins given in Y01602 and
CC   Y11994 to Y12260, respectively. The proteins given represent the signal
CC   peptide and an N-terminal fragment of a secreted protein. The nucleic
CC   acid sequences can be used for producing secreted human gene products.
CC   They can also be used to develop products for diagnosis and therapy.
CC   The proteins obtained may have cytokine activity, cell
CC   proliferation/differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptide can be used
CC   for directing extracellular secretion of a polypeptide or the insertion
CC   of a polypeptide into a membrane, or importing a polypeptide into
CC   a cell.
XX
SQ   Sequence   25 AA;


  Query Match             25.4%;  Score 34;  DB 20;  Length 25;
  Best Local Similarity   50.0%;  Pred. No. 69;
  Matches    9;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy       9 MDFGFPEHLLVDFLQSLS 26
           |  |  |  | :|| |||
Db       1 mgggiaesflcnflvsls 18


RESULT  13
P30225
ID   P30225 standard; Protein; 12 AA.
XX
AC   P30225;
XX
DT   25-MAY-1992  (first entry)
XX
DE   Sequence of interferon (HuIFN) -alpha-61A around amino acid 40.
XX
```

CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 61.9%; Score 83; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 FGFPEHLLVDFLQSLS 26
||||||||||||||||
Db 1 fgfpehllvdflqsls 16

RESULT 10
Y91230
ID Y91230 standard; peptide; 16 AA.
XX
AC Y91230;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:108.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens
CC containing them, are used to induce a T helper cell response,
CC specifically against Plasmodium falciparum, cholesteryl ester transport
CC protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC peptide immunogens may be used for prevention and/or treatment of
CC infections (HIV, foot-and-mouth disease or malaria); for cancer
CC immunotherapy; for inhibition of the action of luteinising hormone
CC releasing hormone (LHRH) for contraception, treatment of hormone-
CC dependent cancer, prevention of boar taint in meat, and
CC immunocastration); for promoting the growth of animals; or for
CC treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC Th (functional in genetically diverse subjects) into an immunogen
CC improves capacity to induce a strong T helper cell-mediated immune
CC response, resulting in production of antibodies against a target antigen.
CC Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC Sequence Y91121 represents a promiscuous T helper epitope from the
CC measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
XX
SQ Sequence 16 AA;

Query Match 59.0%; Score 79; DB 21; Length 16;
Best Local Similarity 93.8%; Pred. No. 5.7e-06;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 88; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the human
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             82.8%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-10;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGFPEHLLVDFLQSLS 26
          ||||||||||||||||||||||
Db      1 lllqmdfgfpehllvdflqsls 22


RESULT   8
Y13809
ID   Y13809 standard; peptide; 22 AA.
XX
AC   Y13809;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Example 1; Page 73; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;


  Query Match             79.9%;  Score 107;  DB 20;  Length 22;
  Best Local Similarity   95.5%;  Pred. No. 4.3e-10;
  Matches   21;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 LLLQMDFGFPEHLLVDFLQSLS 26
          ||||||||||:||||||||||
Db      1 lllqmdfgfpkhllvdflqsls 22


RESULT   9
Y91229
ID   Y91229 standard; peptide; 16 AA.
XX
AC   Y91229;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:107.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13975.
XX
PR   20-JUN-1998;   98US-0100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus  -
XX
PS   Claim 10; Page 50; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
```

```
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
CC   epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC   hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC   synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC   Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC   sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC   antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC   somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC   somatostatin and a Th epitope. Somatostatin immunogens may be used
CC   to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC   antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC   antigenic peptides which may be used to prevent HIV infection of T
CC   cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC   E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC   peptides which may be used in the treatment of allergies. Y91220 is
CC   a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC   protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC   is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC   Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC   used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC   and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC   epitope which may be used to prevent or treat arteriosclerosis and
CC   cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC   B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC   peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC   a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC   an immunostimulatory invasin protein epitope from Yersinia species, and
CC   hinge spacer peptide, both of which may optionally be used in the
CC   antigenic peptides of the invention.
XX
SQ   Sequence   26 AA;

  Query Match             87.3%;  Score 117;  DB 21;  Length 26;
  Best Local Similarity   92.0%;  Pred. No. 1.5e-11;
  Matches   23;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 DGFLLLQMDFGFPEHLLVDFLQSLS 26
           || |||||||||:|||||||||||
Db       2 dgclllqmdfgfpkhllvdflqsls 26


RESULT   6
Y13815
ID   Y13815 standard; peptide; 22 AA.
XX
AC   Y13815;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Disclosure; Page 75; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   22 AA;

  Query Match             82.8%;  Score 111;  DB 20;  Length 22;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-10;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       5 LLLQMDFGFPEHLLVDFLQSLS 26
           ||||||||||||||||||||||
Db       1 lllqmdfgfpehllvdflqsls 22


RESULT   7
Y13821
ID   Y13821 standard; peptide; 22 AA.
XX
AC   Y13821;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Human CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Homo sapiens.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
```

CC epitope. Sequence Y91143 represents a promiscuous Th epitope from
CC hepatitis B virus (HBV) surface antigen, and sequences Y91144-Y91155 are
CC synthetic epitopes derived from this HBV epitope. Y91156-Y91196,
CC Y91227 and Y91242-Y91244 are antigenic peptides comprising an LHRH
CC sequence joined to a promiscuous Th epitope. Y91197 is the LHRH target
CC antigenic peptide used in these LHRH antigenic peptides. Y91200 is
CC somatostatin, and Y91201-Y91207 are antigenic peptides comprising
CC somatostatin and a Th epitope. Somatostatin immunogens may be used
CC to promote growth in livestock. Y91208 is a human CD4 CDR2-like domain
CC antigenic site, and Y91209-Y90211 are MVH Th epitope/CD4 CDR2
CC antigenic peptides which may be used to prevent HIV infection of T
CC cells. Y90212 is a modified version of a human IgE (immunoglobulin
CC E) CH3 domain, and Y90213-Y90219 are Th epitope/IgE CH3 antigenic
CC peptides which may be used in the treatment of allergies. Y91220 is
CC a peptide derived from foot and mouth disease virus (FMDV) VP1 capsid
CC protein and Y91221-Y91222 comprise this peptide and a Th epitope. Y91223
CC is a Plasmodium falciparum circumsporozoite (CS) target antigen, and
CC Y91224-Y91225 comprise the CS antigen and an MVF Th epitope and may be
CC used in a malaria vaccine. Y91228-Y91231 represent CETP-derived peptides
CC and Y91232-Y91241 are immunogens comprising a CETP peptide and a Th
CC epitope which may be used to prevent or treat arteriosclerosis and
CC cardiovascular disease. Y91247 and Y91252-Y91257 are HIV-1 neutralising
CC B-cell epitopes, and Y91248-Y91251 and Y91258-Y91273 are antigenic
CC peptides comprising MVH Th and HIV-1 B-cell epitope which may be used as
CC a component in an anti-HIV-1 vaccine. Y91198 and Y91199 are respectively
CC an immunostimulatory invasin protein epitope from Yersinia species, and
CC hinge spacer peptide, both of which may optionally be used in the
CC antigenic peptides of the invention.
CC Note: Sequence Y91227 is also designated SEQ ID NO:106 in the
CC specification.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 134; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.9e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26
||||||||||||||||||||||||||
Db 1 rdgflllqmdfgfpehllvdflqsls 26

RESULT 4
Y13802
ID Y13802 standard; peptide; 26 AA.
XX
AC Y13802;
XX
DT 08-JUL-1999 (first entry)
XX
DE Rabbit CETP immunogenic fragment.
XX
KW CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW antibody production; cholesteryl ester transfer; therapy;
KW high density lipoprotein; HDL cholesterol concentration;
KW pro-atherogenic dyslipoproteinaemia.
XX
OS Oryctolagus sp.
XX
PN WO9915655-A1.
XX
PD 01-APR-1999.
XX
PF 17-SEP-1998; 98WO-US19366.
XX
PR 19-SEP-1997; 97US-0934367.
XX
PA (MONS) MONSANTO CO.
XX
PI Glenn K, Needleman P;
XX
DR WPI; 1999-276984/23.
XX
PT New recombinant DNA vaccines
XX
PS Claim 15; Page 94; 99pp; English.
XX
CC This sequence represents an immunogenic fragment of the rabbit
CC cholesteryl ester transferase protein (CETP).
CC The invention relates to recombinant DNA vaccines that contain DNA
CC encoding CETP, which can be used for producing antibodies to lessen the
CC transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC method can provide an autogenic immunological process for lessening the
CC transfer of cholesteryl esters from HDL particles and for increasing the
CC HDL cholesterol concentration of a mammal whose blood also contains
CC CETP. The method may be useful in treating human pro-atherogenic
CC dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC method can have an effect that lasts for months as compared to the
CC short-term effects of the small molecule drugs now available.
XX
SQ Sequence 26 AA;

Query Match 87.3%; Score 117; DB 20; Length 26;
Best Local Similarity 92.0%; Pred. No. 1.5e-11;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 DGFLLLQMDFGFPEHLLVDFLQSLS 26
|| |||||||||:|||||||||||
Db 2 dgclllqmdfgfpkhllvdflqsls 26

RESULT 5
Y91231
ID Y91231 standard; peptide; 26 AA.
XX
AC Y91231;
XX
DT 22-MAY-2000 (first entry)
XX
DE Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:109.
XX
KW Promiscuous T-cell epitope; measles virus F protein; MVF;
KW hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS Homo sapiens.
XX
PN WO9966957-A2.
XX
PD 29-DEC-1999.
XX
PF 21-JUN-1999; 99WO-US13975.
XX
PR 20-JUN-1998; 98US-0100412.
XX
PA (UNBI-) UNITED BIOMEDICAL INC.
XX
PI Wang CY;
XX
DR WPI; 2000-160564/14.
XX
PT New artificial T helper cell epitope and derived immunogens with target
PT antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT or human immune deficiency virus -
XX
PS Claim 10; Page 62; 129pp; English.
XX
CC The invention relates to novel promiscuous T helper cell epitopes (Th),
CC and immunogenic peptides comprising the Th epitopes of the invention
CC along with B cell epitopes. The Th epitopes and peptide immunogens

```
XX
CC   A B-cell epitope (W06128) comprising the C-terminal 26 amino acids
CC   of human liver mature cholesteryl ester transfer protein (CETP)
CC   (see also W06127) is involved in a neutral lipid binding or a
CC   transfer activity of CETP.  It can be linked to a universal or
CC   broad range immunogenic T-cell epitope, such as that found at amino
CC   acids 830-843 of tetanus toxoid protein, to produce a synthetic
CC   vaccine (see also W06129) that elicits an immune response against
CC   endogenous CETP activity, thereby treating or preventing
CC   cardiovascular disease, such as atherosclerosis.  It may also be
CC   incorporated into a multivalent vaccine (see also W06131)
CC   including another CETP B-cell epitope.
XX
SQ   Sequence   26 AA;


  Query Match             100.0%;  Score 134;  DB 17;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26
          ||||||||||||||||||||||||||
Db      1 rdgflllqmdfgfpehllvdflqsls 26


RESULT   2
Y13801
ID   Y13801 standard; peptide; 26 AA.
XX
AC   Y13801;
XX
DT   08-JUL-1999  (first entry)
XX
DE   Rabbit CETP immunogenic fragment.
XX
KW   CETP; cholesteryl-ester transfer protein; recombinant DNA vaccine; HDL;
KW   antibody production; cholesteryl ester transfer; therapy;
KW   high density lipoprotein; HDL cholesterol concentration;
KW   pro-atherogenic dyslipoproteinaemia.
XX
OS   Oryctolagus sp.
XX
PN   WO9915655-A1.
XX
PD   01-APR-1999.
XX
PF   17-SEP-1998;   98WO-US19366.
XX
PR   19-SEP-1997;   97US-0934367.
XX
PA   (MONS ) MONSANTO CO.
XX
PI   Glenn K,  Needleman P;
XX
DR   WPI; 1999-276984/23.
XX
PT   New recombinant DNA vaccines
XX
PS   Claim 15; Page 85; 99pp; English.
XX
CC   This sequence represents an immunogenic fragment of the rabbit
CC   cholesteryl ester transferase protein (CETP).
CC   The invention relates to recombinant DNA vaccines that contain DNA
CC   encoding CETP, which can be used for producing antibodies to lessen the
CC   transfer of cholesteryl esters from high density lipoprotein (HDL). The
CC   method can provide an autogenic immunological process for lessening the
CC   transfer of cholesteryl esters from HDL particles and for increasing the
CC   HDL cholesterol concentration of a mammal whose blood also contains
CC   CETP. The method may be useful in treating human pro-atherogenic
CC   dyslipoproteinaemias characterised by low HDL/LDL cholesterol ratios. The
CC   method can have an effect that lasts for months as compared to the
CC   short-term effects of the small molecule drugs now available.
XX
SQ   Sequence   26 AA;


  Query Match             100.0%;  Score 134;  DB 20;  Length 26;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26
          ||||||||||||||||||||||||||
Db      1 rdgflllqmdfgfpehllvdflqsls 26


RESULT   3
Y91228
ID   Y91228 standard; peptide; 26 AA.
XX
AC   Y91228;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Human cholesteryl transport protein (CETP) peptide, SEQ ID NO:106.
XX
KW   Promiscuous T-cell epitope; measles virus F protein; MVF;
KW   hepatitis B virus surface antigen; HBV; immunogenic; B-cell epitope;
KW   luteinising hormone releasing hormone; LHRH; contraceptive; anticancer;
KW   somatostatin; growth promotion; CD4 receptor; HIV-1; antiviral; FMDV;
KW   foot and mouth disease virus; immunoglobulin E; IgE; anti-allergic;
KW   Plasmodium falciparum; circumsporozoite; antimalarial; CETP;
KW   cholesteryl ester transport protein; anti-arteriosclerotic.
XX
OS   Homo sapiens.
XX
PN   WO9966957-A2.
XX
PD   29-DEC-1999.
XX
PF   21-JUN-1999;   99WO-US13975.
XX
PR   20-JUN-1998;   98US-0100412.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PI   Wang CY;
XX
DR   WPI; 2000-160564/14.
XX
PT   New artificial T helper cell epitope and derived immunogens with target
PT   antigenic site, for immunization against e.g. malaria, arteriosclerosis
PT   or human immune deficiency virus
XX
PS   Claim 10; Page 49; 129pp; English.
XX
CC   The invention relates to novel promiscuous T helper cell epitopes (Th),
CC   and immunogenic peptides comprising the Th epitopes of the invention
CC   along with B cell epitopes. The Th epitopes and peptide immunogens
CC   containing them, are used to induce a T helper cell response,
CC   specifically against Plasmodium falciparum, cholesteryl ester transport
CC   protein (CETP) or HIV epitopes, but more generally against any pathogen,
CC   immunoreactive self-antigen or tumour antigen. The Th epitopes and
CC   peptide immunogens may be used for prevention and/or treatment of
CC   infections (HIV, foot-and-mouth disease or malaria); for cancer
CC   immunotherapy; for inhibition of the action of luteinising hormone
CC   releasing hormone (LHRH) for contraception, treatment of hormone-
CC   dependent cancer, prevention of boar taint in meat, and
CC   immunocastration); for promoting the growth of animals; or for
CC   treating allergies or arteriosclerosis. Incorporation of a promiscuous
CC   Th (functional in genetically diverse subjects) into an immunogen
CC   improves capacity to induce a strong T helper cell-mediated immune
CC   response, resulting in production of antibodies against a target antigen.
CC   Th can replace carrier proteins and pathogen-derived T helper epitopes.
CC   Sequence Y91121 represents a promiscuous T helper epitope from the
CC   measles virus F (MVF) protein and sequences Y91122-Y91142, Y91226 and
CC   Y91245-Y91246 represent synthetic Th epitopes based on the MVF Th
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:16 ; Search time 26.24 Seconds
(without alignments)
33.881 Million cell updates/sec

Title: US-08-934-367-29
Perfect score: 134
Sequence: 1 RDGFLLLQMDFGFPEHLLVDFLQSLS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 143201

Minimum DB seq length: 0
Maximum DB seq length: 26

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_36:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	134	100.0	26	17	W06128	Human cholesteryl
2	134	100.0	26	20	Y13801	Rabbit CETP immuno
3	134	100.0	26	21	Y91228	Human cholesteryl
4	117	87.3	26	20	Y13802	Rabbit CETP immuno
5	117	87.3	26	21	Y91231	Human cholesteryl
6	111	82.8	22	20	Y13815	Rabbit CETP immuno
7	111	82.8	22	20	Y13821	Human CETP immunog
8	107	79.9	22	20	Y13809	Rabbit CETP immuno
9	83	61.9	16	21	Y91229	Human cholesteryl
10	79	59.0	16	21	Y91230	Human cholesteryl
11	53	39.6	11	18	W24294	Human/Rabbit CETP
12	34	25.4	25	20	Y12067	Human 5' EST secre
13	33	24.6	12	4	P30225	Sequence of interf
14	33	24.6	12	9	P80053	Sequence of human
15	33	24.6	14	20	Y27814	Human secreted pro
16	33	24.6	21	18	W38080	PPPPY motif contai
17	33	24.6	24	18	W22196	Endogenous TCR alp
18	32	23.9	21	19	W65690	Fibronectin bindin
19	32	23.9	22	20	Y19066	Lecithin:cholester
20	32	23.9	22	20	Y18812	Lecithin:cholester
21	32	23.9	22	20	Y18549	Lecithin:cholester
22	32	23.9	22	20	Y19320	Lecithin:cholester
23	31	23.1	12	20	Y32796	Mammalian prolacti
24	31	23.1	21	19	W65666	Fibronectin bindin
25	31	23.1	22	19	W65669	Peptide #13. Synt
26	30	22.4	17	21	Y68335	Amyotrophic latera
27	30	22.4	20	15	R61276	Transactivating pr
28	30	22.4	20	19	W41190	Tax protein fragme
29	30	22.4	22	17	R95899	Fragment #4 of 7-a
30	29	21.6	10	19	W78506	SH2 domain binding
31	29	21.6	11	20	W99442	Interleukin-2 rece
32	29	21.6	11	21	Y82915	Peptide exhibiting
33	29	21.6	13	18	W10883	MAb anti-HBsAg bin
34	29	21.6	19	16	W13199	Fragment of p53 bi
35	29	21.6	19	19	W57353	Human WBP1 immunog
36	29	21.6	20	16	R72269	Glutamic acid deca
37	29	21.6	20	20	W73610	Human myelin basic
38	29	21.6	20	21	Y85547	Human MBP peptide
39	29	21.6	20	21	Y59570	GAD65 fragment, pe
40	29	21.6	21	21	Y69793	Human interleukin-
41	29	21.6	23	20	Y27486	E. coli beta-subun
42	29	21.6	24	20	Y13053	Human secreted pro
43	29	21.6	25	20	Y42791	Human sulphamidase
44	29	21.6	25	20	W90234	Human sulphamidase
45	28.5	21.3	20	20	Y10894	Amino acid sequenc

ALIGNMENTS

RESULT 1
W06128
ID W06128 standard; Peptide; 26 AA.
XX
AC W06128;
XX
DT 07-FEB-1997 (first entry)
XX
DE Human cholesteryl ester transfer protein C-terminal B-cell epitope.
XX
KW Cholesteryl ester transfer protein; CETP; antigen; vaccine;
KW cardiovascular disease; atherosclerosis; B-cell epitope.
XX
OS Homo sapiens.
XX
PN WO9634888-A1. — not us / declaration by applicant
XX
PD 07-NOV-1996.
XX
PF 01-MAY-1996; 96WO-US06147.
XX
PR 01-MAY-1995; 95US-0432483.
XX
PA (TCEL-) T CELL SCI INC.
XX
PI Rittershaus CW, Thomas LJ;
XX
DR WPI; 1996-506103/50.
XX
PT Cholesteryl ester transfer protein B cell epitope linked to T cell
PT epitope - used to generate vaccine to regulate CETP activity for
PT decreasing the risk of developing a cardiovascular disease e.g.
PT atherosclerosis
XX
PS Claim 5; Page 41; 72pp; English.

```
Db      1 IVGGVKAQA 9


RESULT   7
O04817
ID   O04817      PRELIMINARY;      PRT;    19 AA.
AC   O04817;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   HYPOTHETICAL 2.3 KDA PROTEIN (FRAGMENT).
OS   Sporobolus stapfianus.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Sporobolus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Blomstedt C.K., Gianello R.D., Neale A.D., Hamill J.D., Gaff D.F.;
RL   Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Y10781; CAA71753.1; -.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2270 MW;  4C644625813F244E CRC64;


  Query Match             25.5%;  Score 26;  DB 10;  Length 19;
  Best Local Similarity   27.3%;  Pred. No. 7.8e+02;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 QEIFQEVVGGF 11
          |::: ::  ||
Db      1 QKLYDDLTSGF 11


RESULT   8
Q9R4Z4
ID   Q9R4Z4      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4Z4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AMBIENT-TEMPERATURE FIMBRIA; ATF.
OS   Proteus mirabilis.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Proteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94222573.
RA   Massad G., Bahrani F.K., Mobley H.L.;
RT   "Proteus mirabilis fimbriae: identification, isolation, and
RT   characterization of a new ambient-temperature fimbria.";
RL   Infect. Immun. 62:1989-1994(1994).
SQ   SEQUENCE   20 AA;  2017 MW;  B2FC481003883AE2 CRC64;


  Query Match             25.5%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 8.3e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     10 GFPSQAQVTV 19
          | |:  :|||
Db      4 GTPAPTEVTV 13


RESULT   9
Q9UCA3
ID   Q9UCA3      PRELIMINARY;      PRT;    20 AA.
AC   Q9UCA3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   PROSTASIN (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 94308140.
RA   Yu J.X., Chao L., Chao J.;
RT   "Prostasin is a novel human serine proteinase from seminal fluid.
RT   Purification, tissue distribution, and localization in prostate
RT   gland.";
RL   J. Biol. Chem. 269:18843-18848(1994).
DR   INTERPRO; IPR001254; -.
DR   PFAM; PF00089; trypsin; 1.
SQ   SEQUENCE   20 AA;  2108 MW;  EDF142A1F9F880FE CRC64;


  Query Match             25.5%;  Score 26;  DB 4;  Length 20;
  Best Local Similarity   41.7%;  Pred. No. 8.3e+02;
  Matches    5;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      7 VVGGFPSQAQVT 18
          | | :| | |  :|
Db      8 VAGQWPWQVSIT 19


RESULT  10
Q9S8T0
ID   Q9S8T0      PRELIMINARY;      PRT;    20 AA.
AC   Q9S8T0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   JACALIN BETA-SUBUNIT.
OS   Artocarpus champeden.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Rosales; Moraceae;
OC   Artocarpus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93152601.
RA   Ngoc L.D., Brillard M., Hoebeke J.;
RT   "The alpha- and beta-subunits of the jacalins are cleavage products
RT   from a 17-kDa precursor.";
RL   Biochim. Biophys. Acta 1156:219-222(1993).
SQ   SEQUENCE   20 AA;  2071 MW;  8D5AE975F4D0E212 CRC64;


  Query Match             25.0%;  Score 25.5;  DB 10;  Length 20;
  Best Local Similarity   56.2%;  Pred. No. 1e+03;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    1;  Gaps    1;

Qy      3 IFQEVVGGFPSQAQVT 18
          | | |: | |  ||||
Db      6 ISQTVIVG-PWGAQVT 20


RESULT  11
Q9TWL4
ID   Q9TWL4      PRELIMINARY;      PRT;    18 AA.
AC   Q9TWL4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   25 KDA CHYMOTRYPSIN-LIKE ENZYME (FRAGMENT).
OS   Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Oestroidea; Calliphoridae; Lucilia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95219141.
RA   Casu R.E., Pearson R.D., Jarmey J.M., Cadogan L.C., Riding G.A.,
RA   Tellam R.L.;
RT   "Excretory/secretory chymotrypsin from Lucilia cuprina: purification,
```

```
RT   enzymatic specificity and amino acid sequence deduced from mRNA.";
RL   Insect Mol. Biol. 3:201-211(1994).
SQ   SEQUENCE   18 AA;  1922 MW;  352EB0729B126B11 CRC64;


  Query Match             24.5%;  Score 25;  DB 5;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.1e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       5 QEVVGGFPSQ 14
           :  || || |
Db       6 EATVGQFPYQ 15


RESULT  12
Q06140
ID   Q06140      PRELIMINARY;      PRT;    19 AA.
AC   Q06140;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MAJOR OUTER MEMBRANE PROTEIN P1.16 (PORIN)(FRAGMENT).
GN   PORA.
OS   Neisseria meningitidis.
OC   Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93116032.
RA   Wedege E., Dalseg R., Caugant D.A., Poolman J.T., Froholm L.O.;
RT   "Expression of an inaccessible P1.7 subtype epitope on meningococcal
RT   class 1 proteins.";
RL   J. Med. Microbiol. 38:23-28(1993).
CC   -!- SUBUNIT: MONOMER.
DR   EMBL; S51901; AAB24741.1; -.
KW   Outer membrane; Antigen.
FT   NON_TER       1      1
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1773 MW;  AFAC572B037E0484 CRC64;


  Query Match             24.5%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   42.9%;  Pred. No. 1.2e+03;
  Matches    6;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       5 QEVVGGFPSQAQVT 18
           |   ||   | :||
Db       2 QAANGGASGQVKVT 15


RESULT  13
Q9R5I1
ID   Q9R5I1      PRELIMINARY;      PRT;    19 AA.
AC   Q9R5I1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   NITRILE HYDRATASE ALPHA SUBUNIT (FRAGMENT).
OS   Brevibacterium.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Micrococcineae; Brevibacteriaceae.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92407758.
RA   Duran R., Chion C.K., Bigey F., Arnaud A., Galzy P.;
RT   "The N-terminal amino acid sequences of Brevibacterium sp. R312
RT   nitrile hydratase.";
RL   J. Basic Microbiol. 32:13-19(1992).
SQ   SEQUENCE   19 AA;  1923 MW;  84726D1A1282FB63 CRC64;


  Query Match             24.5%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 PSQAQVT 18
           |:|| ||
Db      13 PAQAPVT 19


RESULT  14
Q9U8F7
ID   Q9U8F7      PRELIMINARY;      PRT;    18 AA.
AC   Q9U8F7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PAIRED DOMAIN TRANSCRIPTIONAL FACTOR EY (FRAGMENT).
GN   EYELESS.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Xu P.-X., Yoon A., Heaney S., Zhang X., Michelson A., Maas R.L.;
RT   "The regulation of Pax6 is conserved between mice and flies.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF089733; AAD54001.1; -.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  1680 MW;  642A9C256071D610 CRC64;


  Query Match             23.5%;  Score 24;  DB 5;  Length 18;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+03;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 VGGFPS 13
           ||| |:
Db      12 VGGIPA 17


RESULT  15
Q62637
ID   Q62637      PRELIMINARY;      PRT;    19 AA.
AC   Q62637;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   XANTHINE DEHYDROGENASE (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE DAWLEY;
RX   MEDLINE; 94268906.
RA   Chow C.W., Clark M., Rinaldo J., Chalkley R.;
RT   "Identification of the rat xanthine dehydrogenase/oxidase promoter.";
RL   Nucleic Acids Res. 22:1846-1854(1994).
DR   EMBL; U08123; AAB60444.1; -.
FT   NON_TER       1      1
FT   VARIANT      19     19       H -> Q.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2128 MW;  A02841D686F50139 CRC64;


  Query Match             23.5%;  Score 24;  DB 11;  Length 19;
  Best Local Similarity   28.6%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 QEIFQEVVGGFPSQ 14
           :|: | |  |   :
Db       4 EELLQSVCAGLAEE 17
```

Search completed: December 21, 2000, 08:35:40
Job time: 603 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:50 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-33
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	31.8	10	3	US-08-159-339A-935	Sequence 935, App
2	31	29.0	20	1	US-08-440-861-11	Sequence 11, Appl
3	30	28.0	10	3	US-08-159-339A-1076	Sequence 1076, Ap
4	30	28.0	14	1	US-08-483-021-3	Sequence 3, Appli
5	30	28.0	20	2	US-08-468-540B-1	Sequence 1, Appli
6	30	28.0	20	2	US-08-606-639A-5	Sequence 5, Appli
7	30	28.0	20	3	US-08-297-395-14	Sequence 14, Appl
8	29	27.1	13	3	US-08-974-775-12	Sequence 12, Appl
9	29	27.1	14	3	US-08-974-775-11	Sequence 11, Appl
10	28	26.2	14	3	US-08-974-775-39	Sequence 39, Appl
11	27	25.2	10	3	US-08-159-339A-558	Sequence 558, App
12	27	25.2	13	1	US-08-221-580-7	Sequence 7, Appli
13	27	25.2	13	4	PCT-US95-04018-69	Sequence 69, Appl
14	27	25.2	15	2	US-08-432-871C-73	Sequence 73, Appl
15	26.5	24.8	20	1	US-07-678-974D-8	Sequence 8, Appli
16	26.5	24.8	20	2	US-08-945-168-13	Sequence 13, Appl
17	26	24.3	8	3	US-08-159-339A-1049	Sequence 1049, Ap
18	26	24.3	8	3	US-08-925-002-49	Sequence 49, Appl
19	26	24.3	11	2	US-08-540-388-1	Sequence 1, Appli
20	26	24.3	12	1	US-07-778-233B-23	Sequence 23, Appl
21	26	24.3	12	1	US-07-963-321-23	Sequence 23, Appl
22	26	24.3	12	1	US-08-290-641-23	Sequence 23, Appl
23	26	24.3	12	1	US-08-548-540-23	Sequence 23, Appl
24	26	24.3	12	4	PCT-US96-09809-23	Sequence 23, Appl
25	26	24.3	13	1	US-08-469-005A-13	Sequence 13, Appl
26	26	24.3	13	1	US-08-188-426-7	Sequence 7, Appli
27	26	24.3	13	2	US-08-469-009-7	Sequence 7, Appli
28	26	24.3	13	2	US-08-469-007-7	Sequence 7, Appli
29	26	24.3	13	3	US-08-974-775-16	Sequence 16, Appl
30	26	24.3	13	3	US-08-462-351-8	Sequence 8, Appli
31	26	24.3	13	5	5194425-9	Patent No. 5194425
32	26	24.3	14	3	US-08-974-775-15	Sequence 15, Appl
33	26	24.3	18	2	US-08-951-871-10	Sequence 10, Appl
34	26	24.3	19	2	US-08-482-142-80	Sequence 80, Appl
35	26	24.3	19	2	US-08-613-235-6	Sequence 6, Appli
36	26	24.3	19	2	US-08-478-572-80	Sequence 80, Appl
37	25.5	23.8	20	2	US-08-107-676-21	Sequence 21, Appl
38	25	23.4	4	5	5425936-4	Patent No. 5425936
39	25	23.4	4	5	5433940-1	Patent No. 5433940
40	25	23.4	5	1	US-08-022-381A-12	Sequence 12, Appl
41	25	23.4	5	1	US-08-475-827A-12	Sequence 12, Appl
42	25	23.4	5	5	5196404-17	Patent No. 5196404
43	25	23.4	5	5	5433940-22	Patent No. 5433940
44	25	23.4	7	1	US-08-482-880-32	Sequence 32, Appl
45	25	23.4	7	2	US-08-273-274-32	Sequence 32, Appl

ALIGNMENTS

```
RESULT 1
US-08-159-339A-935
; Sequence 935, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
;   APPLICANT: Kubo, Ralph T.
;   APPLICANT: Grey, Howard M.
;   APPLICANT: Sette, Alessandro
;   APPLICANT: Celis, Esteban
;   TITLE OF INVENTION: HLA Binding peptides and Their
;   TITLE OF INVENTION: Uses
;   NUMBER OF SEQUENCES: 1254
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Townsend and Townsend and Crew LLP
;     STREET: Two Embarcadero Center, Eighth Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Diskette
;     COMPUTER: IBM Compatible
;     OPERATING SYSTEM: DOS
;     SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/159,339A
;     FILING DATE: 29-NOV-1993
;     CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/926,666
;     FILING DATE: 07-AUG-1992
;     APPLICATION NUMBER: US 08/027,746
;     FILING DATE: 05-MAR-1993
;     APPLICATION NUMBER: US 08/103,396
;     FILING DATE: 06-AUG-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Weber, Ellen Lauver
;     REGISTRATION NUMBER: 32,762
;     REFERENCE/DOCKET NUMBER: 018623-005030US
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (415) 576-0200
;     TELEFAX: (415) 576-0300
;     TELEX:
; INFORMATION FOR SEQ ID NO: 935:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 10 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-159-339A-935
```

```
  Query Match            31.8%;  Score 34;  DB 3;  Length 10;
  Best Local Similarity  50.0%;  Pred. No. 4;
  Matches    5;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 RPDQQHSVAY 18
          :||:|: |:|
Db      1 QPDKQYDVSY 10


RESULT   2
US-08-440-861-11
; Sequence 11, Application US/08440861
; Patent No. 5710126
;  GENERAL INFORMATION:
;    APPLICANT:  Griffith, Irwin J.
;    APPLICANT:  Kuo, Mei-Chang
;    APPLICANT:  Luqman, Mohammad
;    TITLE OF INVENTION:  T CELL EPITOPES OF RYEGRASS POLLEN
;    TITLE OF INVENTION:  ALLERGEN
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII Text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/440,861
;      FILING DATE:  15-MAY-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/106,016
;      FILING DATE:  31-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Amy E. Mandragouras
;      REGISTRATION NUMBER:  36,207
;      REFERENCE/DOCKET NUMBER:  IPC-075 (IMI-040cp)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 227-7400
;      TELEFAX:  (617) 227-5941
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
US-08-440-861-11


  Query Match            29.0%;  Score 31;  DB 1;  Length 20;
  Best Local Similarity  46.2%;  Pred. No. 30;
  Matches    6;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      6 LFPRPDQQHSVAY 18
          | |: |  : |||
Db      3 LIPKLDTAYDVAY 15


RESULT   3
US-08-159-339A-1076
; Sequence 1076, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1076:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-1076


  Query Match            28.0%;  Score 30;  DB 3;  Length 10;
  Best Local Similarity  50.0%;  Pred. No. 20;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      9 RPDQQHSVAY 18
          || |:|    |
Db      1 RPSQRHGSKY 10


RESULT   4
US-08-483-021-3
; Sequence 3, Application US/08483021
; Patent No. 5734023
;  GENERAL INFORMATION:
;    APPLICANT:  Nag, Bishwajit
;    APPLICANT:  Clark, Brian R.
;    APPLICANT:  Sharma, Somesh
;    APPLICANT:  McConnell, Harden
;    TITLE OF INVENTION:  MHC Subunit Conjugates Useful in
;    TITLE OF INVENTION:  Ameliorating Deleterious Immune Responses
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew
;      STREET:  One Market Plaza, Steuart Street Tower
;      CITY:  San Francisco
;      STATE:  California
```

```
;       COUNTRY:  USA
;       ZIP:  94105-1492
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/483,021
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/210,594
;       FILING DATE:  23-JUN-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/367,751
;       FILING DATE:  21-JUN-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/635,840
;       FILING DATE:  12-DEC-1990
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/690,840
;       FILING DATE:  23-APR-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/793,938
;       FILING DATE:  19-NOV-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/978,946
;       FILING DATE:  18-NOV-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/208,072
;       FILING DATE:  03-MAR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/397,533
;       FILING DATE:  02-MAR-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bastian, Kevin L.
;       REGISTRATION NUMBER:  34,774
;       REFERENCE/DOCKET NUMBER:  14058-16-1-4
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 543-9600
;       TELEFAX:  (415) 543-5043
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     FEATURE:
;       NAME/KEY:  Modified-site
;       LOCATION:  1
;       OTHER INFORMATION:  /product= "Xaa = N-acetyl alanine"
;     FEATURE:
;       NAME/KEY:  Peptide
;       LOCATION:  1..14
;       OTHER INFORMATION:  /note= "rat myelin basic protein
;       OTHER INFORMATION:  peptide analog of MBP(1-14)4-A"
US-08-483-021-3


  Query Match             28.0%;  Score 30;  DB 1;  Length 14;
  Best Local Similarity   50.0%;  Pred. No. 29;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   5
US-08-468-540B-1
; Sequence 1, Application US/08468540B
; Patent No. 5858980
;  GENERAL INFORMATION:
;    APPLICANT:  Weiner, Howard
;    APPLICANT:  Hafler, David
;    APPLICANT:  Miller, Ariel
;    APPLICANT:  Al-Sabbagh, Ahmad
;    TITLE OF INVENTION:  SUPPRESSION OF T-CELL PROLIFERATION
;    TITLE OF INVENTION:  USING PEPTIDE FRAGMENTS OF MYELIN BASIC PROTEIN
;    NUMBER OF SEQUENCES:  25
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby P.C.
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/468,540B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jacobs, Seth H
;      REGISTRATION NUMBER:  32,140
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-527-7700
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5858980e
US-08-468-540B-1


  Query Match             28.0%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   6
US-08-606-639A-5
; Sequence 5, Application US/08606639A
; Patent No. 5939400
;  GENERAL INFORMATION:
;    APPLICANT:  Steinman, Lawrence
;    APPLICANT:  Waisman, Ari
;    TITLE OF INVENTION:  DNA VACCINATION FOR INDUCTION OF
;    TITLE OF INVENTION:  SUPPRESSIVE T CELL RESPONSE
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
;      STREET:  4 Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/606,639A
;       FILING DATE:
;       CLASSIFICATION:  514
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Sherwood, Pamela J.
;       REGISTRATION NUMBER:  36,677
;       REFERENCE/DOCKET NUMBER:  A-63005/PJS STAN-187
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 494-8700
;       TELEFAX:  (415) 494-8771
;       TELEX:  910 277299 FHT UR
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-606-639A-5


  Query Match           28.0%;  Score 30;  DB 2;  Length 20;
  Best Local Similarity  50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   7
US-08-297-395-14
; Sequence 14, Application US/08297395A
; Patent No. 6039947
; GENERAL INFORMATION:
;  APPLICANT: Howard L. Weiner
;  APPLICANT: David A. Hafler
;  TITLE OF INVENTION: PEPTIDES DERIVED FROM IMMUNODOMINANT
;  TITLE OF INVENTION: EPITOPES OF MYELIN BASIC PROTEIN
;  FILE REFERENCE: 1010/05723US3
;  CURRENT APPLICATION NUMBER: US/08/297,395A
;  CURRENT FILING DATE: 1994-08-11
;  EARLIER APPLICATION NUMBER: 08/059,189
;  EARLIER FILING DATE: 1993-05-06
;  EARLIER APPLICATION NUMBER: 07/502,559
;  EARLIER FILING DATE: 1990-03-30
;  EARLIER APPLICATION NUMBER: PCT/US88/02139
;  EARLIER FILING DATE: 1988-06-24
;  EARLIER APPLICATION NUMBER: 07/065,734
;  EARLIER FILING DATE: 1987-06-24
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
;   LENGTH: 20
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-08-297-395-14


  Query Match           28.0%;  Score 30;  DB 3;  Length 20;
  Best Local Similarity  50.0%;  Pred. No. 44;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAY 18
           || |:|   |
Db       5 RPSQRHGSKY 14


RESULT   8
US-08-974-775-12
; Sequence 12, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,775
;      FILING DATE:  20-NOV-1997
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  7814/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-775-12


  Query Match           27.1%;  Score 29;  DB 3;  Length 13;
  Best Local Similarity  45.5%;  Pred. No. 40;
  Matches    5;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
           | |   |  ||
Db       3 PYPQGNHEAAY 13


RESULT   9
US-08-974-775-11
; Sequence 11, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/974,775
;       FILING DATE:  20-NOV-1997
;       CLASSIFICATION:  436
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Martin, Alice O.
;       REGISTRATION NUMBER:  35,601
;       REFERENCE/DOCKET NUMBER:  7814/27
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  312-321-4200
;       TELEFAX:  312-321-4299
;   INFORMATION FOR SEQ ID NO:  11:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  14 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-974-775-11


  Query Match             27.1%;  Score 29;  DB 3;  Length 14;
  Best Local Similarity   45.5%;  Pred. No. 44;
  Matches    5;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
           | |   |  ||
Db       3 PYPQGNHEAAY 13


RESULT  10
US-08-974-775-39
; Sequence 39, Application US/08974775
; Patent No. 6096706
;  GENERAL INFORMATION:
;    APPLICANT:  Toback, F. Gary
;    APPLICANT:  Walsh-Reitz, Margaret
;    TITLE OF INVENTION:  GROWTH-PROMOTING PROTEINS AND PEPTIDES
;    TITLE OF INVENTION:  FOR KIDNEY EPITHELIAL CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BRINKS, HOFER, GILSON & LIONE
;      STREET:  NBC Tower - Suite 3600, 455 N. Cityfront
;      STREET:  Plaza Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60611-5599
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/974,775
;      FILING DATE:  20-NOV-1997
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Martin, Alice O.
;      REGISTRATION NUMBER:  35,601
;      REFERENCE/DOCKET NUMBER:  7814/27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-321-4200
;      TELEFAX:  312-321-4299
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-974-775-39


  Query Match             26.2%;  Score 28;  DB 3;  Length 14;
  Best Local Similarity   36.4%;  Pred. No. 66;
  Matches    4;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       8 PRPDQQHSVAY 18
           | |:  |  :|
Db       3 PYPEGNHEASY 13


RESULT  11
US-08-159-339A-558
; Sequence 558, Application US/08159339A
; Patent No. 6037135
;  GENERAL INFORMATION:
;    APPLICANT:  Kubo, Ralph T.
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Celis, Esteban
;    TITLE OF INVENTION:  HLA Binding peptides and Their
;    TITLE OF INVENTION:  Uses
;    NUMBER OF SEQUENCES:  1254
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,339A
;      FILING DATE:  29-NOV-1993
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/926,666
;      FILING DATE:  07-AUG-1992
;      APPLICATION NUMBER:  US 08/027,746
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  558:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-558


  Query Match             25.2%;  Score 27;  DB 3;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 67;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       7 FPRPDQQH 14
           |  ||::|
Db       1 FTTPDKKH 8
```

```
RESULT  12
US-08-221-580-7
; Sequence 7, Application US/08221580
; Patent No. 5519000
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Mervic, Miljenko
;    APPLICANT:  Weber, Robert W.
;    TITLE OF INVENTION:  Tumor Necrosis Factor Inhibitors
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5519000ris
;      STREET:  One Liberty Place 46th Floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  United States of America
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/221,580
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  CCOR-0140
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
;    MOLECULE TYPE:  peptide
US-08-221-580-7


  Query Match             25.2%;  Score 27;  DB 1;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 90;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 QQHSVAYTF 20
           | ||  :||
Db       5 QSHSWPFTF 13


RESULT  13
PCT-US95-04018-69
; Sequence 69, Application PC/TUS9504018
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    APPLICANT:  Mervic, Miljenko
;    APPLICANT:  Weber, Robert W.
;    TITLE OF INVENTION:  Tumor Necrosis Factor Inhibitors
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
;      ADDRESSEE:  Norris
;      STREET:  One Liberty Place  46th Floor
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19403
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/04018
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,580
;      FILING DATE:  01-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,583
;      FILING DATE:  01-APR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/221,581
;      FILING DATE:  01-APR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DeLuca, Mark
;      REGISTRATION NUMBER:  33,229
;      REFERENCE/DOCKET NUMBER:  CCOR-0232
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  both
;    MOLECULE TYPE:  peptide
PCT-US95-04018-69


  Query Match             25.2%;  Score 27;  DB 4;  Length 13;
  Best Local Similarity   55.6%;  Pred. No. 90;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 QQHSVAYTF 20
           | ||  :||
Db       5 QSHSWPFTF 13


RESULT  14
US-08-432-871C-73
; Sequence 73, Application US/08432871C
; Patent No. 5877010
;  GENERAL INFORMATION:
;    APPLICANT:  Loeb, Lawrence A.
;    APPLICANT:  Black, Margaret E.
;    TITLE OF INVENTION:  THYMIDINE KINASE MUTANTS
;    NUMBER OF SEQUENCES:  104
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/432,871C
;      FILING DATE:  02-MAY-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  240052.409C1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
```

```
;       TELEFAX:  (206) 682-6031
;       TELEX:  3723836
;   INFORMATION FOR SEQ ID NO:  73:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-432-871C-73


  Query Match             25.2%;  Score 27;  DB 2;  Length 15;
  Best Local Similarity   42.9%;  Pred. No. 1.1e+02;
  Matches    6;  Conservative    2;  Mismatches    2;  Indels    4;  Gaps    1;

Qy       5 FLFPRPDQQHSVAY 18
           |:| |    | :||
Db       1 FIFDR----HPIAY 10


RESULT  15
US-07-678-974D-8
; Sequence 8, Application US/07678974D
; Patent No. 5629146
;  GENERAL INFORMATION:
;     APPLICANT:  DILLNER, JOAKIM
;     APPLICANT:  DILLNER, LENA
;     TITLE OF INVENTION:  METHOD FOR DETECTION OF HUMAN PAPILLOMAVIRUS
;     TITLE OF INVENTION:
;     NUMBER OF SEQUENCES:  67
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  BERMAN & AISENBERG
;       STREET:  1730 RHODE ISLAND AVENUE, N.W.,
;       CITY:  WASHINGTON
;       STATE:  D.C.
;       COUNTRY:  U.S.A.
;       ZIP:  20036-3186
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/678,974D
;       FILING DATE:  25-JUN-1991
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  AISENBERG, Irwin M.
;       REGISTRATION NUMBER:  19,007
;       REFERENCE/DOCKET NUMBER:  SG19171
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-293-1404
;       TELEFAX:  202-872-0493
;       TELEX:  440 069 AIS UI
;   INFORMATION FOR SEQ ID NO:  8:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-07-678-974D-8


  Query Match             24.8%;  Score 26.5;  DB 1;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 1.8e+02;
  Matches    8;  Conservative    0;  Mismatches    5;  Indels    5;  Gaps    1;

Qy       4 KFLFP-----RPDQQHSV 16
           || ||      || |  |
Db       2 KFGFPDTSFYNPDTQRLV 19


Search completed: December 21, 2000, 08:31:50
Job time: 373 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:07 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-33
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	25.2	12	2	S10059	tachykinin - Afric
2	27	25.2	20	2	JP0053	ribosomal protein
3	26	24.3	20	2	A54083	p190/210, fatty ac
4	25	23.4	14	2	A60770	cell surface alloa
5	25	23.4	15	2	PS0221	gastrin-releasing
6	25	23.4	15	2	A41436	alpha-macroglobuli
7	24	22.4	11	2	JN0023	substance P - chic
8	24	22.4	11	2	S23306	substance P - Atla
9	24	22.4	12	2	S47395	T-cell antigen rec
10	24	22.4	15	2	S29207	avenin gamma-4 - o
11	24	22.4	15	2	A28965	ribulose-bisphosph
12	24	22.4	17	2	B25348	UDPglucose--glycog
13	24	22.4	19	2	S19613	globin - polychaet
14	24	22.4	19	2	S69166	ferredoxin b - Jap
15	24	22.4	20	2	JP0055	ribosomal protein
16	24	22.4	20	2	A34859	helothermine - Mex
17	23	21.5	9	2	A60108	exotoxin A - Strep
18	23	21.5	10	2	S19296	16K protein - poul
19	23	21.5	11	2	S78026	ribosomal protein
20	23	21.5	11	2	S33300	probable substance
21	23	21.5	11	2	S23308	substance P - rain
22	23	21.5	16	2	A60551	leukocyte elastase
23	23	21.5	20	2	A47105	dystroglycan - chi
24	22	20.6	7	4	I56695	hypothetical L2 pr
25	22	20.6	11	2	B60409	kassinin-like pept
26	22	20.6	11	2	C60409	kassinin-like pept
27	22	20.6	11	2	E60409	substance P-like p
28	22	20.6	14	2	S60353	amylopullulanase -
29	22	20.6	18	2	A35704	cytochrome P450 ol
30	22	20.6	18	2	A42016	mammary-derived gr
31	22	20.6	19	2	JP0054	ribosomal protein
32	21	19.6	9	4	S15594	orf 1 rara 5'-regi
33	21	19.6	11	1	SPHO	substance P - hors
34	21	19.6	11	1	A60654	substance P - guin
35	21	19.6	12	2	S47391	T-cell antigen rec
36	21	19.6	15	2	S26515	T-cell receptor al
37	21	19.6	16	2	S57517	T cell receptor be
38	21	19.6	17	2	JH0785	DNA-directed RNA p
39	21	19.6	17	2	S05033	photosystem II pro
40	21	19.6	18	2	S55002	protein 1 - Legion
41	21	19.6	20	2	S18582	hypothetical prote
42	21	19.6	20	2	A56046	urinary tract ston
43	20	18.7	8	2	S66646	cardioacceleratory
44	20	18.7	10	1	XAVI6B	angiotensin-conver
45	20	18.7	10	2	A61007	hementin (EC 3.4.-

ALIGNMENTS

RESULT 1
S10059
tachykinin - African tree frog (Kassina maculata)
N;Alternate names: hylambates-kassinin
C;Species: Kassina maculata
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 24-Mar-1999
C;Accession: S10059
R;Yasuhara, T.; Nakajima, T.; Erspamer, G.F.; Erspamer, V.
Biomed. Res. 2, 613-617, 1981
A;Title: New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and hylambatin, i
A;Reference number: S07436
A;Accession: S10059
A;Molecule type: protein
A;Residues: 1-12 <YAS>
A;Experimental source: skin
A;Note: the source is designated as Hylambates maculatus
C;Keywords: amidated carboxyl end; neuropeptide; tachykinin
F;12/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 25.2%; Score 27; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       8 PRPDQ 12
           |:|||
Db       3 PKPDQ 7
```

RESULT 2
JP0053
ribosomal protein L30 - Bacillus macerans (fragment)
C;Species: Bacillus macerans
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994
C;Accession: JP0053
R;Ochi, K.
submitted to JIPID, February 1994
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal
A;Reference number: JP0042
A;Accession: JP0053
A;Molecule type: protein
A;Residues: 1-20 <OCH>
C;Keywords: protein biosynthesis; ribosome

Query Match 25.2%; Score 27; DB 2; Length 20;
Best Local Similarity 35.7%; Pred. No. 2.5e+02;
Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

```
Qy       2 MVKFLFPRPDQQHS 15
           :|: |  ||: | :
Db       7 LVRSLIGRPEXQRT 20
```

```
RESULT   3
A54083
p190/210, fatty acid synthase, p140exo2 strand exchange protein activator - fission yeas
C;Species: Schizosaccharomyces pombe
C;Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C;Accession: A54083
R;Kaslan, E.; Heyer, W.D.
J. Biol. Chem. 269, 14103-14110, 1994
A;Title: Schizosaccharomyces pombe fatty acid synthase mediates DNA strand exchange in v
A;Reference number: A54083; MUID:94245730
A;Accession: A54083
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <KAS>
A;Note: sequence extracted from NCBI backbone (NCBIP:148744)
C;Superfamily: yeast fatty-acid synthase


  Query Match             24.3%;  Score 26;  DB 2;  Length 20;
  Best Local Similarity   36.4%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
            ||: :  :|:|
Db       2 RPEVEQELAHT 12


RESULT   4
A60770
cell surface alloantigen gp60 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 07-Feb-1997
C;Accession: A60770
R;Schadendorf, D.; Yamaguchi, H.; Old, L.J.; Srivastava, P.K.
J. Immunol. 142, 1621-1625, 1989
A;Title: A novel heteromorphic human cell surface alloantigen, gp60, defined by a human
A;Reference number: A60770; MUID:89140352
A;Accession: A60770
A;Molecule type: protein
A;Residues: 1-14 <SCH>
C;Comment: This protein is an alloantigen in human populations but is not found on cells
C;Comment: This protein exists in both membrane bound and cytosolic forms.
C;Keywords: glycoprotein; polymorphism; surface antigen


  Query Match             23.4%;  Score 25;  DB 2;  Length 14;
  Best Local Similarity   66.7%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      14 HSVAYT 19
           ||||:|
Db       3 HSVSFT 8


RESULT   5
PS0221
gastrin-releasing peptide - laughing frog (fragment)
C;Species: Rana ridibunda (laughing frog)
C;Date: 23-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 03-Mar-1995
C;Accession: PS0221
R;Conlon, J.M.; O'Harte, F.; Vaudry, H.
Biochem. Biophys. Res. Commun. 178, 526-530, 1991
A;Title: Primary structures of the bombesin-like neuropeptides in frog brain show that b
A;Reference number: PQ0177; MUID:91315477
A;Accession: PS0221
A;Molecule type: protein
A;Residues: 1-15 <CON>
A;Experimental source: brain
C;Keywords: neuropeptide


  Query Match             23.4%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   80.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PDQQH 14
           | |||
Db       2 PSQQH 6


RESULT   6
A41436
alpha-macroglobulin - green seaturtle (fragment)
C;Species: Chelonia mydas (green seaturtle)
C;Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 18-Jun-1993
C;Accession: A41436
R;Osada, T.; Sasaki, T.; Ikai, A.
J. Biochem. 103, 212-217, 1988
A;Title: Purification and characterization of alpha-macroglobulin and ovomacroglobuli
A;Reference number: A41436; MUID:88227890
A;Accession: A41436
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <OSA>


  Query Match             23.4%;  Score 25;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 4e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 VMVKFLF 7
           |:| |||
Db       8 VLVPFLF 14


RESULT   7
JN0023
substance P - chicken
C;Species: Gallus gallus (chicken)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jul-1997
C;Accession: JN0023
R;Conlon, J.M.; Katsoulis, S.; Schmidt, W.E.; Thim, L.
Regul. Pept. 20, 171-180, 1988
A;Title: [Arg3]substance P and neurokinin A from chicken small intestine.
A;Reference number: JN0023; MUID:88204263
A;Accession: JN0023
A;Molecule type: protein
A;Residues: 1-11 <CON>
C;Superfamily: substance P precursor
C;Keywords: amidated carboxyl end; tachykinin
F;11/Modified site: amidated carboxyl end (Met) #status predicted


  Query Match             22.4%;  Score 24;  DB 2;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
           ||| |
Db       2 PRPQQ 6


RESULT   8
S23306
substance P - Atlantic cod
C;Species: Gadus morhua (Atlantic cod)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 01-Feb-1999
C;Accession: S23306
R;Jensen, J.; Conlon, J.M.
Eur. J. Biochem. 206, 659-664, 1992
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the
A;Reference number: S23186; MUID:92298992
A;Accession: S23306
A;Molecule type: protein
```

A;Residues: 1-11 <JEN>
A;Experimental source: brain
C;Function:
A;Description: may play a physiological role in the regulation of cardiovascular and gas
A;Note: substance P is derived by post-translational processing of preprotachykinin A
C;Keywords: neuropeptide; amidated carboxyl end; tachykinin
F;11/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 22.4%; Score 24; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      8 PRPDQ 12
          ||| |
Db      2 PRPQQ 6
```

RESULT 9
S47395
T-cell antigen receptor VJ junction beta chain - human
C;Species: Homo sapiens (man)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C;Accession: S47395
R;Lehner, P.J.
submitted to the EMBL Data Library, August 1994
A;Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c
A;Reference number: S47355
A;Accession: S47395
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-12 <LEH>
A;Cross-references: EMBL:Z35715; NID:g527525; PIDN:CAA84784.1; PID:g527526
C;Keywords: T-cell receptor

Query Match 22.4%; Score 24; DB 2; Length 12;
Best Local Similarity 62.5%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy     13 QHSVAYTF 20
          | |  |||
Db      5 QGSYGYTF 12
```

RESULT 10
S29207
avenin gamma-4 - oat (fragment)
N;Alternate names: CIP-1; coeliac immunoreactive protein 1
C;Species: Avena sativa (oat)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C;Accession: S29207
R;Rocher, A.; Colilla, F.; Ortiz, M.L.; Mendez, E.
FEBS Lett. 310, 37-40, 1992
A;Title: Identification of the three major coeliac immunoreactive proteins and one alpha
A;Reference number: S29207; MUID:92405739
A;Accession: S29207
A;Molecule type: protein
A;Residues: 1-15 <ROC>
A;Experimental source: endosperm
C;Superfamily: gliadin
C;Keywords: prolamin; seed

Query Match 22.4%; Score 24; DB 2; Length 15;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      8 PRPDQQ 13
          | |:||
Db      7 PYPEQQ 12
```

RESULT 11
A28965
ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain - spinach (fragments)
C;Species: Spinacia oleracea (spinach)
C;Date: 22-Dec-1988 #sequence_revision 22-Dec-1988 #text_change 23-Feb-1997
C;Accession: A28965
R;Mulligan, R.M.; Houtz, R.L.; Tolbert, N.E.
Proc. Natl. Acad. Sci. U.S.A. 85, 1513-1517, 1988
A;Title: Reaction-intermediate analogue binding by ribulose bisphosphate carboxylase/
cetylated proline.
A;Reference number: A28965; MUID:88144466
A;Accession: A28965
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <MUL>
C;Keywords: carbon-carbon lyase; carboxy-lyase; chloroplast

Query Match 22.4%; Score 24; DB 2; Length 15;
Best Local Similarity 62.5%; Pred. No. 5.9e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      4 KFLFPRPD 11
          || ||  |
Db      6 KFEFPAMD 13
```

RESULT 12
B25348
UDPglucose--glycogen glucosyltransferase (EC 2.4.1.11) P-2 peptide - rabbit (fragment
N;Alternate names: UDPglucose-glycogen glucosyltransferase
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 06-Dec-1996
C;Accession: B25348
R;Wang, Y.; Bell, A.W.; Hermodson, M.A.; Roach, P.J.
J. Biol. Chem. 261, 16909-16915, 1986
A;Title: Liver isozyme of rabbit glycogen synthase. Amino acid sequences surrounding
A;Reference number: A92570; MUID:87057401
A;Accession: B25348
A;Molecule type: protein
A;Residues: 1-17 <WAN>
C;Function:
A;Description: catalyzes the alpha-1,4-glucosylation of glycogen by UDPglucose produc
C;Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 22.4%; Score 24; DB 2; Length 17;
Best Local Similarity 60.0%; Pred. No. 6.7e+02;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy      6 LFPRP 10
          ::|||
Db      1 MYPRP 5
```

RESULT 13
S19613
globin - polychaete (Eudistylia vancouveri) (fragment)
C;Species: Eudistylia vancouveri
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 05-Dec-1998
C;Accession: S19613
R;Qabar, A.N.; Stern, M.S.; Walz, D.A.; Chiu, J.T.; Timkovich, R.; Wall, J.S.; Kapp,
J. Mol. Biol. 222, 1109-1129, 1991
A;Title: Hierarchy of globin complexes. The quaternary structure of the extracellular
A;Reference number: S19532; MUID:92106333
A;Accession: S19613
A;Molecule type: protein
A;Residues: 1-19 <QAB>
A;Experimental source: plume
C;Complex: dodecamers, each consisting of a trimer of tetramers of globin chains; dod
C;Keywords: calcium; dodecamer; heme; homotetramer; oxygen carrier

Query Match 22.4%; Score 24; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 7.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      12 QQHSVAY 18
           || |:||
Db      12 QQWSMAY 18
```

RESULT 14
S69166
ferredoxin b - Japanese radish (fragments)
C;Species: Kaiware daikon (Japanese radish)
C;Date: 10-Mar-1998 #sequence_revision 10-Mar-1998 #text_change 17-Apr-1998
C;Accession: S69166
R;Obata, S.; Nishimura, M.; Nagai, K.; Sakihama, N.; Shin, M.
Arch. Biochem. Biophys. 316, 797-802, 1995
A;Title: Four ferredoxins from Japanese radish leaves.
A;Reference number: S69164; MUID:95168867
A;Accession: S69166
A;Molecule type: protein
A;Residues: 1-19 <OBA>
C;Keywords: 2Fe-2S; electron transfer; iron-sulfur protein

Query Match 22.4%; Score 24; DB 2; Length 19;
Best Local Similarity 28.6%; Pred. No. 7.6e+02;
Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
Qy       3 VKFLFPRPDQQHSV 16
           |||: |   ::  :
Db       5 VKFITPETHKEEDM 18
```

RESULT 15
JP0055
ribosomal protein L30 - Bacillus polymyxa (fragment)
C;Species: Bacillus polymyxa
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 28-Oct-1994
C;Accession: JP0055
R;Ochi, K.
submitted to JIPID, February 1994
A;Description: Phylogenetic diversity in the genus Bacillus and comparative ribosomal pr
A;Reference number: JP0042
A;Accession: JP0055
A;Molecule type: protein
A;Residues: 1-20 <OCH>
C;Keywords: protein biosynthesis; ribosome

Query Match 22.4%; Score 24; DB 2; Length 20;
Best Local Similarity 35.7%; Pred. No. 8.1e+02;
Matches 5; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```
Qy       2 MVKFLFPRPDQQHS 15
           :|: |  ||  | :
Db       7 LVRSLIGRPGTQRT 20
```

Search completed: December 21, 2000, 08:30:08
Job time: 272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:32:59 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-33
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1086

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27	25.2	12	1	TKN1_KASMA	P08613 kassina mac
2	25	23.4	15	1	ALB2_TRASC	P81189 trachemys s
3	24	22.4	11	1	TKNA_CHICK	P19850 gallus gall
4	24	22.4	11	1	TKNA_GADMO	P28498 gadus morhu
5	24	22.4	18	1	GALS_SALTY	P41030 salmonella
6	24	22.4	20	1	HELT_HELHO	P46693 heloderma h
7	23	21.5	11	1	TKNA_ONCMY	P28499 oncorhynchu
8	23	21.5	11	1	TKNA_SCYCA	P41333 scyliorhinu
9	22	20.6	11	1	TKN1_PSEGU	P42986 pseudophryn
10	22	20.6	11	1	TKN2_PSEGU	P42987 pseudophryn
11	22	20.6	11	1	TKN4_PSEGU	P42989 pseudophryn
12	22	20.6	18	1	CPAX_BOVIN	P22779 bos taurus
13	21	19.6	9	1	UPA3_HUMAN	P30089 homo sapien
14	21	19.6	11	1	TKNA_HORSE	P01290 equus cabal
15	21	19.6	13	1	ECDE_LYMDI	P80941 lymantria d
16	21	19.6	15	1	MK2A_PALPR	P80409 palomena pr
17	21	19.6	16	1	MK3_PALPR	P80411 palomena pr
18	21	19.6	17	1	PSBL_SYNVU	P12241 synechococc
19	20	18.7	9	1	FAR9_ASCSU	P43172 ascaris suu
20	20	18.7	10	1	BPP2_BOTJA	P01022 bothrops ja
21	20	18.7	11	1	TKNA_RANCA	P22688 rana catesb
22	20	18.7	13	1	BPP1_BOTJA	P01020 bothrops ja
23	20	18.7	13	1	GER1_HORVU	P28525 hordeum vul
24	20	18.7	17	1	SRY_URSAR	P36396 ursus arcto
25	20	18.7	19	1	OXLA_OPHHA	P81383 ophiophagus
26	20	18.7	20	1	ACPH_BOVIN	P80227 bos taurus
27	20	18.7	20	1	LEC3_ARTIN	P18673 artocarpus
28	19	17.8	5	1	PAP2_PARMA	P81864 pardachirus
29	19	17.8	10	1	FARP_LOCMI	P38553 locusta mig
30	19	17.8	11	1	BPP3_BOTIN	P30423 bothrops in
31	19	17.8	11	1	BPP4_BOTIN	P30424 bothrops in
32	19	17.8	11	1	BPPB_AGKHA	P01021 agkistrodon
33	19	17.8	12	1	TKN2_KASMA	P08614 kassina mac
34	19	17.8	12	1	TKN_KASSE	P08611 kassina sen
35	19	17.8	13	1	GER2_HORVU	P28526 hordeum vul
36	19	17.8	13	1	YCIA_SALTY	P25944 salmonella
37	19	17.8	15	1	MK1_PALPR	P80408 palomena pr
38	19	17.8	16	1	MK2B_PALPR	P80410 palomena pr
39	19	17.8	17	1	APID_BOMPA	P81464 bombus pasc
40	19	17.8	17	1	YALA_TRYBB	P17961 trypanosoma
41	19	17.8	20	1	ALAT_PIG	P13191 sus scrofa
42	19	17.8	20	1	PYRR_PYRAP	P37362 pyrrhocoris
43	18.5	17.3	18	1	HEMH_THETS	P80155 theromyzon
44	18	16.8	10	1	UPA2_HUMAN	P30088 homo sapien
45	18	16.8	13	1	CRBL_ICASP	P17237 icaria sp.

ALIGNMENTS

RESULT 1
TKN1_KASMA
ID TKN1_KASMA STANDARD; PRT; 12 AA.
AC P08613;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYLAMBATES KASSININ (GLU(2)-PRO(5) KASSININ).
OS Kassina maculata (African rhacophorid frog) (Hylambates maculatus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;
OC Kassina.
RN [1]
RP SEQUENCE.
RA Yasuhara T., Nakajima T., Erspamer G.F., Erspamer V.;
RT "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and
RT hylambatin, in the skin of the African rhacophorid frog Hylambates
RT maculatus.";
RL Biomed. Res. 2:613-617(1981).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR PIR; S10059; S10059.
DR INTERPRO; IPR002040; -.
DR PROSITE; PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1376 MW; 3E756D279DD6DAB7 CRC64;

Query Match 25.2%; Score 27; DB 1; Length 12;
Best Local Similarity 80.0%; Pred. No. 69;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       8 PRPDQ 12
           |:|||
Db       3 PKPDQ 7
```

RESULT 2
ALB2_TRASC
ID ALB2_TRASC STANDARD; PRT; 15 AA.
AC P81189;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE 68 KDA SERUM ALBUMIN (ALB-2) (FRAGMENT).
OS Trachemys scripta (Red-eared slider turtle) (Pseudemys scripta).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Testudines; Cryptodira; Testudinoidea; Emydidae; Trachemys.
RN [1]
RP SEQUENCE.
RX MEDLINE; 98103404.
RA Brown M.A., Chambers G.K., Licht P.;

```
RT   "Purification and partial amino acid sequences of two distinct
RT   albumins from turtle plasma.";
RL   Comp. Biochem. Physiol. 118B:367-374(1997).
CC   -!- FUNCTION: SERUM ALBUMIN, THE MAIN PROTEIN OF PLASMA, HAS A GOOD
CC       BINDING CAPACITY FOR WATER, CA++, NA+, K+, FATTY ACIDS, HORMONES,
CC       BILIRUBIN AND DRUGS. ITS MAIN FUNCTION IS THE REGULATION OF THE
CC       COLLOIDAL OSMOTIC PRESSURE OF BLOOD.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: PLASMA.
CC   -!- MISCELLANEOUS: IN THE RED-EARED SLIDER TURTLE, THERE ARE TWO FORMS
CC       OF ALBUMIN, ALB-1 AND ALB-2.
CC   -!- SIMILARITY: BELONGS TO THE ALB/AFP/VDB FAMILY.
DR   INTERPRO; IPR000264; -.
DR   PROSITE; PS00212; ALBUMIN; PARTIAL.
KW   Plasma; Metal-binding; Lipid-binding; Albumin.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1733 MW;  4B7422B89FF73223 CRC64;


  Query Match             23.4%;  Score 25;  DB 1;  Length 15;
  Best Local Similarity   30.0%;  Pred. No. 1.9e+02;
  Matches    3;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 DQQHSVAYTF 20
           |: |:: : |
Db       3 DETHTLGHXF 12


RESULT   3
TKNA_CHICK
ID   TKNA_CHICK     STANDARD;      PRT;    11 AA.
AC   P19850;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=INTESTINE;
RX   MEDLINE; 88204263.
RA   Conlon J.M., Katsoulis S., Schmidt W.E., Thim L.;
RT   "[Arg3]substance P and neurokinin A from chicken small intestine.";
RL   Regul. Pept. 20:171-180(1988).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; JN0023; JN0023.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1377 MW;  21487FE3C9D6C6C7 CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
            ||| |
Db       2 PRPQQ 6


RESULT   4
TKNA_GADMO
ID   TKNA_GADMO     STANDARD;      PRT;    11 AA.
AC   P28498;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Gadus morhua (Atlantic cod).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidei; Gadidae;
OC   Gadus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 92298992.
RA   Jensen J., Conlon J.M.;
RT   "Substance-P-related and neurokinin-A-related peptides from the brain
RT   of the cod and trout.";
RL   Eur. J. Biochem. 206:659-664(1992).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION (BY SIMILARITY).
SQ   SEQUENCE   11 AA;  1315 MW;  214860D759D6C6C7 CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
            ||| |
Db       2 PRPQQ 6


RESULT   5
GALS_SALTY
ID   GALS_SALTY     STANDARD;      PRT;    18 AA.
AC   P41030;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   MGL REPRESSOR AND GALACTOSE ULTRAINDUCTION FACTOR (FRAGMENT).
GN   GALS.
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Salmonella.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=LT2;
RX   MEDLINE; 89112167.
RA   Benner-Luger D., Boos W.;
RT   "The mglB sequence of Salmonella typhimurium LT2; promoter analysis
RT   by gene fusions and evidence for a divergently oriented gene coding
RT   for the mgl repressor.";
RL   Mol. Gen. Genet. 214:579-587(1988).
CC   -!- FUNCTION: REPRESSOR OF THE MGL OPERON. BINDS GALACTOSE AND
CC       D-FUCOSE AS INDUCERS. GALS BINDS TO AN OPERATOR DNA SEQUENCE
CC       WITHIN ITS OWN CODING SEQUENCE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LACI FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   STYGENE; SG10521; GALS.
DR   INTERPRO; IPR000843; -.
DR   PROSITE; PS00356; HTH_LACI_FAMILY; PARTIAL.
KW   Transcription regulation; DNA-binding; Repressor.
FT   NON_TER      1      1
SQ   SEQUENCE   18 AA;  2133 MW;  74A77150FA49FB8B CRC64;


  Query Match             22.4%;  Score 24;  DB 1;  Length 18;
```

```
  Best Local Similarity   57.1%;  Pred. No. 3.5e+02;
  Matches    4;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      12 QQHSVAY 18
           ::||||:
Db       3 RRHSVAW 9


RESULT   6
HELT_HELHO
ID   HELT_HELHO     STANDARD;      PRT;     20 AA.
AC   P46693;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   HELOTHERMINE (FRAGMENT).
OS   Heloderma horridum horridum (Mexican beaded lizard).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Anguimorpha; Helodermatidae;
OC   Heloderma.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 90260878.
RA   Mochca-Morales J., Martin B.M., Possani L.D.;
RT   "Isolation and characterization of helothermine, a novel toxin from
RT   Heloderma horridum horridum (Mexican beaded lizard) venom.";
RL   Toxicon 28:299-309(1990).
CC   -!- FUNCTION: TOXIC TO MICE; INDUCES LETHARGY, PARTIAL PARALYSIS OF
CC       REAR LIMBS AND LOWERING OF BODY TEMPERATURE, SUGGESTING THAT IT
CC       MIGHT BE A HYPOTHERMIC TOXIN.
CC   -!- MISCELLANEOUS: THE COMPLETE PROTEIN HAS AN APPARENT MW OF 25 KDA
CC       AND A PI OF 6.8.
DR   PIR; A34859; A34859.
KW   Toxin.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2156 MW;  91D62B36F7B4F940 CRC64;


  Query Match            22.4%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 4e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 PDQQHSV 16
           ||||  :
Db      14 PDQQTEI 20


RESULT   7
TKNA_ONCMY
ID   TKNA_ONCMY     STANDARD;      PRT;     11 AA.
AC   P28499;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 92298992.
RA   Jensen J., Conlon J.M.;
RT   "Substance-P-related and neurokinin-A-related peptides from the brain
RT   of the cod and trout.";
RL   Eur. J. Biochem. 206:659-664(1992).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S23307; S23307.
DR   PIR; S23308; S23308.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION (BY SIMILARITY).
SQ   SEQUENCE   11 AA;  1358 MW;  214860DEC9D6D1F7 CRC64;


  Query Match            21.5%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
           ||| |
Db       2 PRPHQ 6


RESULT   8
TKNA_SCYCA
ID   TKNA_SCYCA     STANDARD;      PRT;     11 AA.
AC   P41333;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 93292508.
RA   Waugh D., Wang Y., Hazon N., Balment R.J., Conlon J.M.;
RT   "Primary structures and biological activities of substance-P-related
RT   peptides from the brain of the dogfish, Scyliorhinus canicula.";
RL   Eur. J. Biochem. 214:469-474(1993).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; S33300; S33300.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1278 MW;  214860DEC9D6D867 CRC64;


  Query Match            21.5%;  Score 23;  DB 1;  Length 11;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 PRPDQ 12
           ||| |
Db       2 PRPGQ 6


RESULT   9
TKN1_PSEGU
ID   TKN1_PSEGU     STANDARD;      PRT;     11 AA.
AC   P42986;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   KASSININ-LIKE PEPTIDE K-I (PG-KI).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
```

```
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; B60409; B60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1269 MW;  3DBA7C37C9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PHPDE 6


RESULT  10
TKN2_PSEGU
ID   TKN2_PSEGU     STANDARD;      PRT;    11 AA.
AC   P42987;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   KASSININ-LIKE PEPTIDE K-II (PG-KII).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; C60409; C60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1246 MW;  3A247C37C9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PNPDE 6


RESULT  11
TKN4_PSEGU
ID   TKN4_PSEGU     STANDARD;      PRT;    11 AA.
AC   P42989;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P-LIKE PEPTIDE I (PG-SPI).
OS   Pseudophryne guentheri (Frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Pseudophryne.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 90287814.
RA   Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,
RA   Roberts J.D., Melchiorri P., Erspamer V.;
RT   "Six novel tachykinin- and bombesin-related peptides from the skin of
RT   the Australian frog Pseudophryne guntheri.";
RL   Peptides 11:299-304(1990).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; E60409; E60409.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1294 MW;  3A247C2CC9CB1AB7 CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          | ||:
Db      2 PNPDE 6


RESULT  12
CPAX_BOVIN
ID   CPAX_BOVIN     STANDARD;      PRT;    18 AA.
AC   P22779;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CYTOCHROME P450 2A (OLF2) (OLFACTIVE) (P52) (EC 1.14.14.1) (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 91027757.
RA   Lazard D., Tal N., Rubinstein M., Khen M., Lancet D., Zupko K.;
RT   "Identification and biochemical analysis of novel olfactory-specific
RT   cytochrome P-450IIA and UDP-glucuronosyl transferase.";
RL   Biochemistry 29:7433-7440(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY
CC       OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY
CC       ACIDS, AND XENOBIOTICS.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
```

```
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A35704; A35704.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum; Olfaction.
FT   NON_TER       1      1
FT   VARIANT       6      6       G -> D.
FT   VARIANT      11     11       A -> E.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2058 MW;  F80746F76CCD77FF CRC64;


  Query Match             20.6%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      8 PRPDQQ 13
          | | ||
Db      5 PGPQQQ 10


RESULT  13
UPA3_HUMAN
ID   UPA3_HUMAN     STANDARD;      PRT;     9 AA.
AC   P30089;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   01-APR-1993 (Rel. 25, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 11) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PLASMA;
RX   MEDLINE; 93092937.
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 4.6, ITS MW IS: 46 KDA.
DR   SWISS-2DPAGE; P30089; HUMAN.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1056 MW;  26F2B1BAF769C737 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 LFPRPD 11
          |||  |
Db      3 LFPXTD 8


RESULT  14
TKNA_HORSE
ID   TKNA_HORSE     STANDARD;      PRT;    11 AA.
AC   P01290;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   SUBSTANCE P.
GN   TAC1 OR NKNA OR TAC2 OR NKA.
OS   Equus caballus (Horse), and Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=HORSE;
RA   Studer R.O., Trzeciak A., Lergier W.;
RT   "Isolation and amino-acid sequence of substance P from horse
RT   intestine.";
RL   Helv. Chim. Acta 56:860-866(1973).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=C.PORCELLUS;
RX   MEDLINE; 90044685.
RA   Murphy R.;
RT   "Primary amino acid sequence of guinea-pig substance P.";
RL   Neuropeptides 14:105-110(1989).
CC   -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC       EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC       SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC       MUSCLES.
CC   -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR   PIR; A01558; SPHO.
DR   PIR; A60654; A60654.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; 1.
KW   Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1349 MW;  3E757FE3C9D6C6C7 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 11;
  Best Local Similarity   60.0%;  Pred. No. 6.8e+02;
  Matches    3;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 PRPDQ 12
          |:| |
Db      2 PKPQQ 6


RESULT  15
ECDE_LYMDI
ID   ECDE_LYMDI     STANDARD;      PRT;    13 AA.
AC   P80941;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   TESTIS ECDYSIOTROPIN PEPTIDE E (TE).
OS   Lymantria dispar (Gypsy moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Lymantriidae; Lymantria.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=BRAIN;
RX   MEDLINE; 97387807.
RA   Loeb M.J., Wagner R.M., Woods C.W., Gelman D.G., Harrison D.,
RA   Bell R.A.;
RT   "Naturally occurring analogs of Lymantria testis ecdysiotropin, a
RT   gonadotropin isolated from brains of Lymantria dispar pupae.";
RL   Arch. Insect Biochem. Physiol. 36:37-50(1997).
CC   -!- FUNCTION: STIMULATES SYNTHESIS OF ECDYSTEROID IN THE TESTES
CC       OF LARVAE AND PUPAE.
SQ   SEQUENCE   13 AA;  1357 MW;  1841B4CA3275B764 CRC64;


  Query Match             19.6%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      8 PRPDQQHS 15
          | ||   |
Db      5 PNPDTPDS 12


Search completed: December 21, 2000, 08:33:00
Job time: 443 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:40 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-33
Perfect score: 107
Sequence: 1 VMVKFLFPRPDQQHSVAYTF 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34	31.8	17	3	Q06800	Q06800 saccharomyc
2	32	29.9	18	12	Q73487	Q73487 human immun
3	30	28.0	19	12	Q9W9U1	Q9w9u1 simian herp
4	30	28.0	20	12	Q9WLD2	Q9wld2 simian herp
5	29	27.1	18	12	Q73461	Q73461 human immun
6	29	27.1	18	12	Q73463	Q73463 human immun
7	29	27.1	18	12	Q73465	Q73465 human immun
8	29	27.1	18	12	Q73467	Q73467 human immun
9	29	27.1	18	12	Q73469	Q73469 human immun
10	29	27.1	18	12	Q73471	Q73471 human immun
11	29	27.1	18	12	Q73473	Q73473 human immun
12	29	27.1	18	12	Q73475	Q73475 human immun
13	29	27.1	18	12	Q73477	Q73477 human immun
14	29	27.1	18	12	Q73479	Q73479 human immun
15	29	27.1	18	12	Q73481	Q73481 human immun
16	29	27.1	18	12	Q73483	Q73483 human immun
17	29	27.1	18	12	Q73485	Q73485 human immun
18	29	27.1	18	12	Q73489	Q73489 human immun
19	29	27.1	18	12	Q73491	Q73491 human immun
20	29	27.1	18	12	Q73493	Q73493 human immun
21	29	27.1	18	12	Q73495	Q73495 human immun
22	29	27.1	18	12	Q73497	Q73497 human immun
23	29	27.1	18	12	Q73499	Q73499 human immun
24	29	27.1	18	12	Q73501	Q73501 human immun
25	29	27.1	18	12	Q73503	Q73503 human immun
26	29	27.1	18	12	Q73505	Q73505 human immun
27	29	27.1	18	12	Q73507	Q73507 human immun
28	29	27.1	18	12	Q73509	Q73509 human immun
29	29	27.1	18	12	Q73511	Q73511 human immun
30	29	27.1	18	12	Q73513	Q73513 human immun
31	29	27.1	18	12	Q73515	Q73515 human immun
32	29	27.1	18	12	Q73519	Q73519 human immun
33	29	27.1	18	12	Q73521	Q73521 human immun
34	29	27.1	18	12	Q73525	Q73525 human immun
35	29	27.1	18	12	Q73527	Q73527 human immun
36	29	27.1	18	12	Q73529	Q73529 human immun
37	28	26.2	14	12	Q89818	Q89818 murine minu
38	28	26.2	16	12	Q79458	Q79458 human immun
39	28	26.2	19	12	O90625	O90625 simian herp
40	27	25.2	17	12	Q73531	Q73531 human immun
41	27	25.2	17	12	Q73533	Q73533 human immun
42	27	25.2	17	12	Q73535	Q73535 human immun
43	27	25.2	17	12	Q73537	Q73537 human immun
44	27	25.2	17	12	Q73539	Q73539 human immun
45	27	25.2	17	12	Q73541	Q73541 human immun

ALIGNMENTS

RESULT 1
Q06800
ID Q06800 PRELIMINARY; PRT; 17 AA.
AC Q06800;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE N1780.
GN N1780.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96287653.
RA Nasr F., Becam A.M., Herbert C.J.;
RT "The sequence of 36.8 kb from the left arm of chromosome XIV reveals
RT 24 complete open reading frames: 18 correspond to new genes, one of
RT which encodes a protein similar to the human myotonic dystrophy
RT kinase.";
RL Yeast 12:169-175(1996).
DR EMBL; X92517; CAA63292.1; -.
SQ SEQUENCE 17 AA; 2139 MW; BD7E9AAFADF754AF CRC64;

Query Match 31.8%; Score 34; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy      12 QQHSVAYT 19
           ||||| :|
Db       3 QQHSVCFT 10
```

RESULT 2
Q73487
ID Q73487 PRELIMINARY; PRT; 18 AA.
AC Q73487;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE REVERSE TRANSCRIPTASE (FRAGMENT).

```
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45053; AAB04204.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2191 MW;  B36910F4583BA0C3 CRC64;


  Query Match             29.9%;  Score 32;  DB 12;  Length 18;
  Best Local Similarity   38.5%;  Pred. No. 63;
  Matches    5;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           ::|: |  ||::|
Db       2 LLKWGFTTPDKKH 14


RESULT   3
Q9W9U1
ID   Q9W9U1      PRELIMINARY;      PRT;    19 AA.
AC   Q9W9U1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   GLYCOPROTEIN G (FRAGMENT).
GN   US4.
OS   Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=20620, 9400371, 7709642, 7709609;
RA   Eberle R., Smith A., Black D.;
RT   "Molecular evidence for distinct genotypes of monkey B virus
RT   (Herpesvirus simiae) which are related to the Macaque host species.";
RL   J. Virol. 0:0-0(1998).
DR   EMBL; AF082804; AAC34447.1; -.
DR   EMBL; AF082812; AAC34111.1; -.
DR   EMBL; AF082813; AAC34114.1; -.
DR   EMBL; AF082814; AAC34117.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2220 MW;  E585178F064B278B CRC64;


  Query Match             28.0%;  Score 30;  DB 12;  Length 19;
  Best Local Similarity   54.5%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
           |   || | ||
Db       2 RSPSQHRVRYT 12


RESULT   4
Q9WLD2
ID   Q9WLD2      PRELIMINARY;      PRT;    20 AA.
AC   Q9WLD2;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   GLYCOPROTEIN G (FRAGMENT).
GN   US4.
OS   Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Alphaherpesvirinae; Simplexvirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=12930;
RA   Eberle R., Smith A., Black D.;
RT   "Molecular evidence for distinct genotypes of monkey B virus
RT   (Herpesvirus simiae) which are related to the Macaque host species.";
RL   J. Virol. 0:0-0(1998).
DR   EMBL; AF082808; AAC34459.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   20 AA;  2377 MW;  E58517E7DBAB278B CRC64;


  Query Match             28.0%;  Score 30;  DB 12;  Length 20;
  Best Local Similarity   54.5%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       9 RPDQQHSVAYT 19
           |   || | ||
Db       3 RSPSQHRVRYT 13


RESULT   5
Q73461
ID   Q73461      PRELIMINARY;      PRT;    18 AA.
AC   Q73461;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45027; AAB04178.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36918D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT   6
Q73463
ID   Q73463      PRELIMINARY;      PRT;    18 AA.
AC   Q73463;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
```

```
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45029; AAB04180.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT   7
Q73465
ID   Q73465      PRELIMINARY;      PRT;    18 AA.
AC   Q73465;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45031; AAB04182.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT   8
Q73467
ID   Q73467      PRELIMINARY;      PRT;    18 AA.
AC   Q73467;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45033; AAB04184.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT   9
Q73469
ID   Q73469      PRELIMINARY;      PRT;    18 AA.
AC   Q73469;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45035; AAB04186.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;
```

```
  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT  10
Q73471
ID   Q73471      PRELIMINARY;      PRT;    18 AA.
AC   Q73471;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45037; AAB04188.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT  11
Q73473
ID   Q73473      PRELIMINARY;      PRT;    18 AA.
AC   Q73473;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45039; AAB04190.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT  12
Q73475
ID   Q73475      PRELIMINARY;      PRT;    18 AA.
AC   Q73475;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45041; AAB04192.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36138D4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 MVKFLFPRPDQQH 14
           :::: |  ||::|
Db       2 LLRWGFTTPDKKH 14


RESULT  13
Q73477
ID   Q73477      PRELIMINARY;      PRT;    18 AA.
AC   Q73477;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
```

```
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45043; AAB04194.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  14
Q73479
ID   Q73479      PRELIMINARY;      PRT;    18 AA.
AC   Q73479;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45045; AAB04196.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


RESULT  15
Q73481
ID   Q73481      PRELIMINARY;      PRT;    18 AA.
AC   Q73481;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   REVERSE TRANSCRIPTASE (FRAGMENT).
GN   POL.
OS   Human immunodeficiency virus type 1.
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RX   MEDLINE; 96242958.
RA   Cleland A., Watson H.G., Robertson P., Ludlam C.A., Brown A.J.;
RT   "Evolution of zidovudine resistance-associated genotypes in human
RT   immunodeficiency virus type 1-infected patients.";
RL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12:6-18(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=PATIENT P74;
RA   Leigh Brown A.J.;
RL   Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U45047; AAB04198.1; -.
KW   RNA-directed DNA polymerase.
FT   NON_TER       1      1
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2219 MW;  B36910F4583BBFFE CRC64;


  Query Match             27.1%;  Score 29;  DB 12;  Length 18;
  Best Local Similarity   30.8%;  Pred. No. 2e+02;
  Matches    4;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 MVKFLFPRPDQQH 14
          :::: |  ||::|
Db      2 LLRWGFTTPDKKH 14


Search completed: December 21, 2000, 08:35:40
Job time: 603 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:36:43 ; Search time 25.87 Seconds
(without alignments)
14.254 Million cell updates/sec

Title: US-08-934-367-34
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 92835

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	28.8	22	3	US-08-940-095-133	Sequence 133, App
2	32	28.8	22	3	US-08-940-093-133	Sequence 133, App
3	32	28.8	22	3	US-08-940-096-133	Sequence 133, App
4	30	27.0	20	1	US-08-199-508-2	Sequence 2, Appli
5	29	26.1	19	1	US-08-399-696-85	Sequence 85, Appl
6	29	26.1	20	2	US-08-484-530-22	Sequence 22, Appl
7	29	26.1	20	2	US-08-827-618A-22	Sequence 22, Appl
8	29	26.1	20	3	US-08-483-952A-22	Sequence 22, Appl
9	28	25.2	10	3	US-08-159-339A-401	Sequence 401, App
10	28	25.2	22	3	US-08-940-095-134	Sequence 134, App
11	28	25.2	22	3	US-08-940-093-134	Sequence 134, App
12	28	25.2	22	3	US-08-940-096-134	Sequence 134, App
13	27	24.3	9	2	US-08-765-783A-83	Sequence 83, Appl
14	27	24.3	15	3	US-08-596-257A-10	Sequence 10, Appl
15	27	24.3	15	3	US-08-860-339-10	Sequence 10, Appl
16	27	24.3	20	1	US-07-678-974D-8	Sequence 8, Appli
17	27	24.3	20	2	US-08-162-149-10	Sequence 10, Appl
18	27	24.3	20	2	US-08-945-168-13	Sequence 13, Appl
19	27	24.3	22	2	US-08-559-524A-12	Sequence 12, Appl
20	27	24.3	22	3	US-08-749-707-12	Sequence 12, Appl
21	26	23.4	7	1	US-07-923-724-47	Sequence 47, Appl
22	26	23.4	7	2	US-08-609-426A-47	Sequence 47, Appl
23	26	23.4	7	2	US-08-374-652C-37	Sequence 37, Appl
24	26	23.4	11	1	US-08-039-778B-6	Sequence 6, Appli
25	26	23.4	12	1	US-08-555-394-13	Sequence 13, Appl
26	26	23.4	12	3	US-08-745-892-13	Sequence 13, Appl
27	26	23.4	13	1	US-08-620-213-3	Sequence 3, Appli
28	26	23.4	16	1	US-08-346-455B-52	Sequence 52, Appl
29	26	23.4	16	3	US-08-977-221-52	Sequence 52, Appl
30	26	23.4	16	4	PCT-US95-06613-52	Sequence 52, Appl
31	26	23.4	17	1	US-08-620-213-1	Sequence 1, Appli
32	26	23.4	19	1	US-08-238-163-20	Sequence 20, Appl
33	26	23.4	19	1	US-07-977-696C-36	Sequence 36, Appl
34	26	23.4	19	1	US-08-129-930B-36	Sequence 36, Appl
35	26	23.4	22	1	US-08-039-778B-5	Sequence 5, Appli
36	25	22.5	6	1	US-08-487-006-76	Sequence 76, Appl
37	25	22.5	6	2	US-08-488-659A-76	Sequence 76, Appl
38	25	22.5	8	3	US-08-582-776C-29	Sequence 29, Appl
39	25	22.5	8	3	US-08-434-831B-29	Sequence 29, Appl
40	25	22.5	10	2	US-08-318-157B-25	Sequence 25, Appl
41	25	22.5	11	1	US-08-039-778B-3	Sequence 3, Appli
42	25	22.5	13	3	US-09-140-149-21	Sequence 21, Appl
43	25	22.5	15	1	US-07-918-181A-17	Sequence 17, Appl
44	25	22.5	15	1	US-07-918-181A-21	Sequence 21, Appl
45	25	22.5	15	1	US-08-080-073-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-08-940-095-133
; Sequence 133, Application US/08940095
; Patent No. 6004925
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    APPLICANT:  Dufourcq, Jean
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,095
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0004-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6004925e
US-08-940-095-133
```

```
  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   2
US-08-940-093-133
; Sequence 133, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,093
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0006-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6037323e
US-08-940-093-133


  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   3
US-08-940-096-133
; Sequence 133, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,096
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0005-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  133:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-133


  Query Match             28.8%;  Score 32;  DB 3;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 33;
  Matches    7;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | ||:|:
Db       7 ERLLEDLLQALN 18


RESULT   4
US-08-199-508-2
; Sequence 2, Application US/08199508
; Patent No. 5717058
;  GENERAL INFORMATION:
;    APPLICANT:  Matthews, Maura-Ann H.
;    APPLICANT:  Stetler, Gary L.
;    APPLICANT:  Anthony-Cahill, Spencer J.
;    APPLICANT:  Anderson, David C.
;    TITLE OF INVENTION:  Modulators of Gene Expression
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Somatogen, Inc.
;      STREET:  5797 Central Avenue
;      CITY:  Boulder
;      STATE:  Colorado
;      ZIP:  80301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.4 Mb storage
```

```
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  System 7.0.1
;       SOFTWARE:  Microsoft Word 5.0a
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/199,508
;       FILING DATE:  February 18, 1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/021,536
;       FILING DATE:  February 23, 1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 5717058ak, Henry P.
;       REGISTRATION NUMBER:  33200
;       REFERENCE/DOCKET NUMBER:  121 CIP
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  303-541-3322
;       TELEFAX:  303-444-3013
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  20
;       TYPE:  amino acid
;       TOPOLOGY:  unknown to applicant
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  yes
US-08-199-508-2


  Query Match             27.0%;  Score 30;  DB 1;  Length 20;
  Best Local Similarity   38.5%;  Pred. No. 63;
  Matches    5;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFLQ 19
           ||:|  ::   | :|
Db       1 FGYPVYVFGDCVQ 13


RESULT   5
US-08-399-696-85
; Sequence 85, Application US/08399696
; Patent No. 5756669
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  p53-BINDING POLYPEPTIDES AND
;    TITLE OF INVENTION:  POLYNUCLEOTIDES ENCODING SAME
;    NUMBER OF SEQUENCES:  126
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/399,696
;      FILING DATE:  02-MAR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/156,671
;      FILING DATE:  22-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15522-000710
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  85:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  19 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-399-696-85


  Query Match             26.1%;  Score 29;  DB 1;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 87;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFL 18
           |||::| |
Db      12 EHLIIDGL 19


RESULT   6
US-08-484-530-22
; Sequence 22, Application US/08484530
; Patent No. 5846740
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisoc
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,530
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Trecartin, Richard F.
;      REGISTRATION NUMBER:  31,801
;      REFERENCE/DOCKET NUMBER:  A-60780-8/RFT/MTK
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-781-1989
;      TELEFAX:  415-398-3249
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-530-22


  Query Match             26.1%;  Score 29;  DB 2;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 93;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           :|| :|  ||
Db       9 IDFHYPNELL 18


RESULT   7
US-08-827-618A-22
; Sequence 22, Application US/08827618A
; Patent No. 5998366
;  GENERAL INFORMATION:
;    APPLICANT:  Tobin, Allan J
;    APPLICANT:  Erlander, Mark G
;    APPLICANT:  Kaufman, Daniel L.
;    TITLE OF INVENTION:  Cloned Glutamic Acid Decarboxylase
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,618A
; FILING DATE: 09-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,725
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/716,909
; FILING DATE: 18-JUN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/586,536
; FILING DATE: 21-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-60780-12/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-827-618A-22

Query Match 26.1%; Score 29; DB 2; Length 20;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 MDFGFPEHLL 14
:|| :| ||
Db 9 IDFHYPNELL 18

RESULT 8
US-08-483-952A-22
; Sequence 22, Application US/08483952A
; Patent No. 6011139
; GENERAL INFORMATION:
; APPLICANT: Tobin, Allan J
; APPLICANT: Erlander, Mark G
; APPLICANT: Kaufman, Daniel L.
; TITLE OF INVENTION: Cloned Glutamic Acid Decarboxylase
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,952A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/123,859
; FILING DATE: 17-SEP-1993
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/716,909
; FILING DATE: 18-JUN-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/586,536
; FILING DATE: 21-SEP-1990
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-60780-7/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-952A-22

Query Match 26.1%; Score 29; DB 3; Length 20;
Best Local Similarity 50.0%; Pred. No. 93;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 5 MDFGFPEHLL 14
:|| :| ||
Db 9 IDFHYPNELL 18

RESULT 9
US-08-159-339A-401
; Sequence 401, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Esteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 1254
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,339A
; FILING DATE: 29-NOV-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/926,666
; FILING DATE: 07-AUG-1992
; APPLICATION NUMBER: US 08/027,746

```
;      FILING DATE:  05-MAR-1993
;      APPLICATION NUMBER:  US 08/103,396
;      FILING DATE:  06-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Weber, Ellen Lauver
;      REGISTRATION NUMBER:  32,762
;      REFERENCE/DOCKET NUMBER:  018623-005030US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  401:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-339A-401


  Query Match             25.2%;  Score 28;  DB 3;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 61;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 FPEHLL 14
             :||||:
Db       2 YPEHLV 7


RESULT  10
US-08-940-095-134
; Sequence 134, Application US/08940095
; Patent No. 6004925
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    APPLICANT:  Dufourcq, Jean
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,095
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0004-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6004925e
US-08-940-095-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | |::|:
Db       7 ERLLEDLLKALN 18


RESULT  11
US-08-940-093-134
; Sequence 134, Application US/08940093
; Patent No. 6037323
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,093
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0006-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6037323e
US-08-940-093-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | |::|:
```

```
Db      7 ERLLEDLLKALN 18


RESULT  12
US-08-940-096-134
; Sequence 134, Application US/08940096
; Patent No. 6046166
;  GENERAL INFORMATION:
;    APPLICANT:  Dasseux, Jean-Louis
;    APPLICANT:  Sekul, Renate
;    APPLICANT:  Buttner, Klaus
;    APPLICANT:  Cornut, Isabelle
;    APPLICANT:  Metz, Gunther
;    TITLE OF INVENTION:  APOLIPOPROTEIN A-I AGONISTS
;    TITLE OF INVENTION:  AND THEIR USE TO TREAT DYSLIPIDEMIC DISORDERS
;    NUMBER OF SEQUENCES:  258
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds LLP
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10036-2811
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/940,096
;      FILING DATE:  29-SEP-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  009196-0005-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  134:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6046166e
US-08-940-096-134


  Query Match             25.2%;  Score 28;  DB 3;  Length 22;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    6;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQSLS 22
           | || | |::|:
Db       7 ERLLEDLLKALN 18


RESULT  13
US-08-765-783A-83
; Sequence 83, Application US/08765783A
; Patent No. 5994524
;  GENERAL INFORMATION:
;    APPLICANT:  Matsushima, Kouji
;    APPLICANT:  Matsumoto, Yoshihiro
;    APPLICANT:  Yamada, Yoshiki
;    APPLICANT:  Sato, Koh
;    APPLICANT:  Tsuchiya, Masayuki
;    APPLICANT:  Yamazaki, Tatsumi
;    TITLE OF INVENTION:  Reshaped Human Antibody to
;    TITLE OF INVENTION:  Interleukin-8
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  2000 Pennsylvania Avenue, NW, suite 5500
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20006-1888
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/765,783A
;      FILING DATE:  07-MAR-1997
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Murashige, Kate H
;      REGISTRATION NUMBER:  29,959
;      REFERENCE/DOCKET NUMBER:  35029-20001.20
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-887-1500
;      TELEFAX:  202-822-0168
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  83:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-765-783A-83


  Query Match             24.3%;  Score 27;  DB 2;  Length 9;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+05;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 QMDFGFP 10
            |  ||||
Db       1 QHHFGFP 7


RESULT  14
US-08-596-257A-10
; Sequence 10, Application US/08596257A
; Patent No. 6001628
;  GENERAL INFORMATION:
;    APPLICANT:  KOSSMAN, Jens
;    APPLICANT:  EMMERMANN, Michael
;    APPLICANT:  VIRGIN, Ivar
;    TITLE OF INVENTION:  DEBRANCHING ENZYMES AND DNA SEQUENCES
;    TITLE OF INVENTION:  CODING THEM, SUITABLE FOR CHANGING THE DEGREE OF BRANCHING
;    TITLE OF INVENTION:  OF AMYLOPECTIN STARCH IN PLANTS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ostrolenk, Faber, Gerb & Soffen
;      STREET:  1180 Avenue of the Americas
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  10036-8403
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/08/596,257A
;       FILING DATE:  08-FEB-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  DE 4327165.0
;       FILING DATE:  09-AUG-1993
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meilman, Edward A.
;       REGISTRATION NUMBER:  24,735
;       REFERENCE/DOCKET NUMBER:  P/951-118
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 382-0700
;       TELEFAX:  (212) 382-0888
;       TELEX:  236925
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     ORIGINAL SOURCE:
;       ORGANISM:  solanum tuberosum
US-08-596-257A-10


  Query Match             24.3%;  Score 27;  DB 3;  Length 15;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPEHLL 14
           | |  ::|  |: ::
Db       1 LEQTNYGLPQQVI 13


RESULT  15
US-08-860-339-10
; Sequence 10, Application US/08860339
; Patent No. 6117665
; GENERAL INFORMATION:
;  APPLICANT: Kossmann, Jens
;  APPLICANT: Emmermann, Michael
;  APPLICANT: Virgin, Ivar
;  APPLICANT: Renz, Andreas
;  TITLE OF INVENTION: DNA MOLECULES CODING FOR DEBRANCHING ENZYMES DERIVED
;  TITLE OF INVENTION: FROM PLANTS
;  FILE REFERENCE: AGREVO-6
;  CURRENT APPLICATION NUMBER: US/08/860,339
;  CURRENT FILING DATE: 1997-11-25
;  EARLIER APPLICATION NUMBER: DE P4447387.7
;  EARLIER FILING DATE: 1994-12-22
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Solanum tuberosum
US-08-860-339-10


  Query Match             24.3%;  Score 27;  DB 3;  Length 15;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+02;
  Matches    4;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLQMDFGFPEHLL 14
           | |  ::|  |: ::
Db       1 LEQTNYGLPQQVI 13


Search completed: December 21, 2000, 08:36:43
Job time: 387 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:13 ; Search time 26.84 Seconds
(without alignments)
52.017 Million cell updates/sec

Title: US-08-934-367-34
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 4315

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.2	19	2	A39504	octamer-binding pr
2	27	24.3	21	2	A32521	hexokinase (EC 2.7
3	26	23.4	15	2	S59492	formate dehydrogen
4	25	22.5	15	2	S71306	heat shock protein
5	25	22.5	16	2	T44936	calmodulin kinase
6	24.5	22.1	14	2	B61597	cytochrome P450 AL
7	24	21.6	13	2	S03879	6-phosphofructokin
8	24	21.6	14	2	A01250	angiotensin precur
9	24	21.6	15	2	A60834	angiotensin I prec
10	24	21.6	17	2	B31769	T-cell receptor de
11	24	21.6	18	2	PN0175	glutathione transf
12	24	21.6	18	2	H75063	hypothetical prote
13	24	21.6	20	2	A37984	ADP,ATP carrier pr
14	23	20.7	7	2	S36662	dermorphin (Lys-7)
15	23	20.7	14	2	PA0015	seed storage prote
16	23	20.7	19	2	I46654	T-cell receptor de
17	23	20.7	21	2	T07683	proteinase inhibit
18	22	19.8	7	1	A61324	dermorphin - Rohde
19	22	19.8	11	2	PT0250	Ig heavy chain CRD
20	22	19.8	12	2	C36201	1-aminocyclopropan
21	22	19.8	14	2	A61002	photosystem II oxy
22	22	19.8	20	2	S72501	protein kinase C i
23	22	19.8	20	2	A31049	calsequestrin, fas
24	22	19.8	20	2	S77989	cytochrome-c oxida
25	22	19.8	20	2	A47105	dystroglycan - chi
26	22	19.8	20	2	S50203	zona pellucida gly
27	22	19.8	21	2	PX0078	alanine dehydrogen
28	21	18.9	9	2	B45796	dihydrolipoamide S
29	21	18.9	9	2	B20569	serum amyloid P-co
30	21	18.9	14	2	PA0104	protein QF200070 -
31	21	18.9	15	2	PH1788	T cell receptor al
32	21	18.9	15	2	S72432	epoxypropan isomer
33	21	18.9	17	2	A60743	ornithine carbamoy
34	21	18.9	18	2	I49408	cytochrome-c oxida
35	21	18.9	19	2	B60822	cytochrome P450 UT
36	21	18.9	19	2	D32071	T-cell receptor de
37	21	18.9	20	2	A60822	cytochrome P450 PB
38	21	18.9	20	2	S07232	ribulose-bisphosph
39	21	18.9	20	2	H49034	nuclear antigen EB
40	21	18.9	20	2	S35921	T-cell receptor ga
41	21	18.9	20	2	A54077	cytochrome b558 -
42	21	18.9	22	2	B54397	ubiquitin-carrier
43	21	18.9	22	2	D56978	collagen alpha 1(I
44	20	18.0	9	2	I46023	growth hormone rec
45	20	18.0	12	2	S56122	type I DNA methylt

ALIGNMENTS

RESULT 1
A39504
octamer-binding protein, Ku-like, 72K chain - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Dec-1991 #sequence_revision 30-Dec-1991 #text_change 30-Sep-1993
C;Accession: A39504
R;May, G.; Sutton, C.; Gould, H.
J. Biol. Chem. 266, 3052-3059, 1991
A;Title: Purification and characterization of Ku-2, an octamer-binding protein relate
A;Reference number: A39504; MUID:91131605
A;Accession: A39504
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <MAY>

Query Match 25.2%; Score 28; DB 2; Length 19;
Best Local Similarity 71.4%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 LQMDFGF 9
           |:|| ||
Db      11 LEMDVGF 17
```

RESULT 2
A32521
hexokinase (EC 2.7.1.1) I peptide III - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 12-Apr-1995
C;Accession: A32521
R;Schirch, D.M.; Wilson, J.E.
Arch. Biochem. Biophys. 257, 1-12, 1987
A;Title: Rat brain hexokinase: amino acid sequence at the substrate hexose binding si
A;Reference number: A90080; MUID:87324917
A;Accession: A32521
A;Molecule type: protein
A;Residues: 1-21 <SCH>
C;Superfamily: human hexokinase I; hexokinase homology
C;Keywords: ATP; glycolysis; phosphotransferase

Query Match 24.3%; Score 27; DB 2; Length 21;
Best Local Similarity 50.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       3 LQMDFGFPEH 12
           |   | || |
Db       3 LGFTFSFPXH 12
```

RESULT 3

S59492
formate dehydrogenase alpha chain - Alcaligenes eutrophus (fragment)
C;Species: Alcaligenes eutrophus
C;Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 07-May-1999
C;Accession: S59492
R;Friedebold, J.; Mayer, F.; Bill, E.; Trautwein, A.X.; Bowien, B.
Biol. Chem. Hoppe-Seyler 376, 561-568, 1995
A;Title: Structural and immunological studies on the soluble formate dehydrogenase from
A;Reference number: S59492; MUID:96145736
A;Accession: S59492
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <FRI>

Query Match 23.4%; Score 26; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 3.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 QMDFGFP 10
::||| |
Db 6 EIDFGTP 12

RESULT 4
S71306
heat shock protein 90 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S71306
R;Conconi, M.; Szweda, L.I.; Levine, R.L.; Stadtman, E.R.; Friguet, B.
Arch. Biochem. Biophys. 331, 232-240, 1996
A;Title: Age-related decline of rat liver multicatalytic proteinase activity and protect
A;Reference number: S71306; MUID:96299287
A;Accession: S71306
A;Molecule type: protein
A;Residues: 1-15 <CON>
A;Experimental source: liver
C;Keywords: heat shock; phosphoprotein; stress-induced protein

Query Match 22.5%; Score 25; DB 2; Length 15;
Best Local Similarity 40.0%; Pred. No. 5.1e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 10 PEHLLVDFLQ 19
|:| :|: |:
Db 6 PDHPIVETLR 15

RESULT 5
T44936
calmodulin kinase 2 - fission yeast (Schizosaccharomyces pombe) (fragment)
C;Species: Schizosaccharomyces pombe
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C;Accession: T44936
R;Alemany, V.; Aligue, R.
submitted to the EMBL Data Library, May 1996
A;Reference number: Z22873
A;Accession: T44936
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-16 <ALE>
A;Cross-references: EMBL:U57982; PIDN:AAD09466.1

Query Match 22.5%; Score 25; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 FGFP 10
||||
Db 5 FGFP 8

RESULT 6
B61597
cytochrome P450 AL-2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C;Accession: B61597
R;Shimeno, H.; Toda, A.; Ogata, S.; Nagamatsu, A.
Drug Metab. Dispos. 19, 291-297, 1991
A;Title: Purification and aminopyrine monooxygenase activity of liver microsomal cyto
A;Reference number: A61597; MUID:91292910
A;Accession: B61597
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-14 <SHI>

Query Match 22.1%; Score 24.5; DB 2; Length 14;
Best Local Similarity 47.1%; Pred. No. 5.7e+02;
Matches 8; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Qy 1 LLLQMDFGFPEHLLVDF 17
||| : | ||| |
Db 3 LLLSLSF-----LLVGF 14

RESULT 7
S03879
6-phosphofructokinase (EC 2.7.1.11) B - rabbit (fragment)
N;Alternate names: phosphofructo-1-kinase B
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 28-Apr-1993
C;Accession: S03879
R;Valaitis, A.P.; Foe, L.G.; Kwiatkowska, D.; Latshaw, S.P.; Kemp, R.G.
Biochim. Biophys. Acta 995, 187-194, 1989
A;Title: The sites of phosphorylation of rabbit brain phosphofructo-1-kinase by cycli
A;Reference number: S03878; MUID:89194250
A;Accession: S03879
A;Molecule type: protein
A;Residues: 1-13 <VAL>
A;Note: the sequence from the summary is inconsistent with that from table I and Fig.
C;Keywords: glycolysis; phosphotransferase

Query Match 21.6%; Score 24; DB 2; Length 13;
Best Local Similarity 71.4%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LQMDFGF 9
| || ||
Db 7 LSMDKGF 13

RESULT 8
A01250
angiotensin precursor - horse (fragment)
C;Species: Equus caballus (domestic horse)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 20-Mar-1998
C;Accession: A92775; A01250
R;Skeggs Jr., L.T.; Kahn, J.R.; Lentz, K.; Shumway, N.P.
J. Exp. Med. 106, 439-453, 1957
A;Reference number: A92775
A;Accession: A92775
A;Molecule type: protein
A;Residues: 1-14 <SKE>
C;Superfamily: antithrombin III
C;Keywords: blood pressure control; hormone; vasoconstrictor
F;1-10/Product: angiotensin I #status experimental <AN1>
F;1-8/Product: angiotensin II #status experimental <AN2>

Query Match 21.6%; Score 24; DB 2; Length 14;

```
  Best Local Similarity   83.3%;  Pred. No. 6.9e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PEHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT   9
A60834
angiotensin I precursor - dog (fragment)
N;Alternate names: angiotensinogen I
N;Contains: angiotensin I
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 13-Mar-1998
C;Accession: A60834
R;Oliver, J.A.
Hypertension 11, 21-27, 1988
A;Title: Purification and partial characterization of canine angiotensinogen.
A;Reference number: A60834; MUID:88113996
A;Accession: A60834
A;Molecule type: protein
A;Residues: 1-15 <OLI>
C;Superfamily: antithrombin III
C;Keywords: glycoprotein; plasma
F;1-10/Product: angiotensin I #status predicted <MAT>


  Query Match             21.6%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   83.3%;  Pred. No. 7.4e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 PEHLLV 15
           | ||||
Db       7 PFHLLV 12


RESULT  10
B31769
T-cell receptor delta-2 chain J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
C;Accession: B31769
R;Loh, E.Y.; Cwirla, S.; Serafini, A.T.; Phillips, J.H.; Lanier, L.L.
Proc. Natl. Acad. Sci. U.S.A. 85, 9714-9718, 1988
A;Title: Human T-cell-receptor delta chain: genomic organization, diversity, and express
A;Reference number: A94221; MUID:89071766
A;Accession: B31769
A;Molecule type: DNA
A;Residues: 1-17 <LOH>
A;Cross-references: GB:L36386; NID:g540455; PIDN:AAA61108.1; PID:g540456
C;Keywords: T-cell receptor


  Query Match             21.6%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   37.5%;  Pred. No. 8.6e+02;
  Matches    6;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPEHLLVD 16
           |  |: ||    |:|:
Db       1 LTAQLFFGKGTQLIVE 16


RESULT  11
PN0175
glutathione transferase (EC 2.5.1.18) 1 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 22-Dec-1995 #sequence_revision 08-Feb-1996 #text_change 15-Jun-1996
C;Accession: PN0175
R;Tsugita, A.; Kamo, M.; Kawakami, M.; Ohki, Y.
submitted to JIPID, December 1995
A;Description: Two dimensional electrophoresis of plant proteins and standardization of
A;Reference number: PN0173
A;Accession: PN0175
A;Molecule type: protein
A;Residues: 1-18 <TSU>
A;Experimental source: leaf
C;Keywords: transferase


  Query Match             21.6%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   80.0%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 DFGFP 10
           ||| |
Db       5 DFGHP 9


RESULT  12
H75063
hypothetical protein PAB7382 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: H75063
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome s
A;Reference number: A75001
A;Accession: H75063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <KAW>
A;Cross-references: GB:AJ248287; GB:AL096836; NID:g5458657; PIDN:CAB50405.1; PID:e151
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB7382


  Query Match             21.6%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   35.7%;  Pred. No. 9.2e+02;
  Matches    5;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPEHLLVD 16
           |::: |  : :|||
Db       3 LRVETGCIDRVLVD 16


RESULT  13
A37984
ADP,ATP carrier protein - yeast (Candida parapsilosis) (fragment)
C;Species: Candida parapsilosis
C;Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 07-Jun-1996
C;Accession: A37984
R;Guerin, B.; Bukusoglu, C.; Rakotomanana, F.; Wohlrab, H.
J. Biol. Chem. 265, 19736-19741, 1990
A;Title: Mitochondrial phosphate transport. N-ethylmaleimide insensitivity correlates
A;Reference number: A37984; MUID:91060585
A;Accession: A37984
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <GUE>


  Query Match             21.6%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 HLLVDFL 18
           : |:|||
Db       5 NFLIDFL 11


RESULT  14
S36662
dermorphin (Lys-7) [validated] - two-colored leaf frog
```

C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 10-Dec-1993 #sequence_revision 13-Mar-1997 #text_change 31-Mar-2000
C;Accession: S36662
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Kreil, G.
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of th
A;Reference number: S21152; MUID:92339502
A;Accession: S36662
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>

Query Match 20.7%; Score 23; DB 2; Length 7;
Best Local Similarity 60.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```
Qy       7 FGFPE 11
           ||:|:
Db       3 FGYPK 7
```

RESULT 15
PA0015
seed storage protein 12S 2 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 11-Jul-1997
C;Accession: PA0015
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensiona
A;Reference number: PA0001
A;Accession: PA0015
A;Molecule type: protein
A;Residues: 1-14 <KAM>
A;Experimental source: seed
C;Keywords: pyroglutamic acid; seed; storage protein
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 20.7%; Score 23; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 5; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       4 QMDFGFPEHL 13
           |   | || |
Db       1 QQYLGVPEQL 10
```

Search completed: December 21, 2000, 08:37:13
Job time: 317 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:38:16 ; Search time 18.39 Seconds
(without alignments)
38.219 Million cell updates/sec

Title: US-08-934-367-34
Perfect score: 111
Sequence: 1 LLLQMDFGFPPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1213

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	24	21.6	14	1	ANGT_HORSE	P01016	equus cabal
2	22	19.8	20	1	CAQS_RAT	P19633	rattus norv
3	22	19.8	20	1	COXN_THUOB	P80980	thunnus obe
4	21	18.9	9	1	SAMP_MUSCA	P19095	mustelus ca
5	20	18.0	11	1	CH60_DROME	P35380	drosophila
6	20	18.0	11	1	TEML_RANTE	P56923	rana tempor
7	20	18.0	14	1	DCMM_PSECF	P19914	pseudomonas
8	20	18.0	16	1	MDH_SYNY4	P80460	synechocyst
9	20	18.0	21	1	ATPB_PHYPA	P80658	physcomitre
10	19.5	17.6	20	1	DPP4_BOVIN	P81425	bos taurus
11	19	17.1	19	1	OXLA_OPHHA	P81383	ophiophagus
12	19	17.1	19	1	UP21_UPEIN	P82027	uperoleia i
13	19	17.1	19	1	UP25_UPEIN	P82031	uperoleia i
14	19	17.1	20	1	SUCB_CANFA	P99507	canis famil
15	19	17.1	21	1	DCMS_PSECA	P19921	pseudomonas
16	19	17.1	21	1	YD90_HAEIN	P45194	haemophilus
17	18	16.2	7	1	ALL7_CYDPO	P82158	cydia pomon
18	18	16.2	8	1	ALL5_CALVO	P41841	calliphora
19	18	16.2	8	1	ALL5_CYDPO	P82156	cydia pomon
20	18	16.2	12	1	GRAR_RANRU	P40754	rana rugosa
21	18	16.2	13	1	ORCK_ORCLI	P37086	orconectes
22	18	16.2	13	1	PSBP_PINPS	P81668	pinus pinas
23	18	16.2	15	1	LMA2_LOCMI	P38497	locusta mig
24	18	16.2	17	1	FLA2_BARBA	P35634	bartonella
25	18	16.2	18	1	RS4_PSEAE	Q52759	pseudomonas
26	18	16.2	19	1	HBB2_UROHA	P18992	uromastyx h
27	18	16.2	19	1	NUO6_SOLTU	P80729	solanum tub
28	18	16.2	20	1	THIO_CANFA	P99505	canis famil
29	18	16.2	20	1	UN05_PINPS	P81674	pinus pinas
30	18	16.2	21	1	CSPS_STRTR	P81622	streptococc
31	18	16.2	21	1	OMP4_PASHA	P80228	pasteurella
32	17	15.3	8	1	CCKN_MACEU	P30369	macropus eu
33	17	15.3	9	1	FAR1_CALVO	P41856	calliphora
34	17	15.3	10	1	CAER_LITXA	P56264	litoria xan
35	17	15.3	12	1	OPS3_DROVI	P17645	drosophila
36	17	15.3	12	1	PSP3_PHYPA	P80662	physcomitre
37	17	15.3	13	1	SODM_CANFA	P54712	canis famil
38	17	15.3	14	1	HCYA_MEGCR	Q10583	megathura c
39	17	15.3	15	1	LPF_ECOLI	P03057	escherichia
40	17	15.3	16	1	LPK1_LOCMI	P20404	locusta mig
41	17	15.3	17	1	GAST_MACMU	P33714	macaca mula
42	17	15.3	19	1	IRBP_CAVPO	P12666	cavia porce
43	17	15.3	19	1	UP24_UPEIN	P82030	uperoleia i
44	17	15.3	20	1	CPA7_PAPSP	P80055	papio sp. (
45	17	15.3	20	1	LPP3_HUMAN	P56643	homo sapien

ALIGNMENTS

RESULT 1
ANGT_HORSE
ID ANGT_HORSE STANDARD; PRT; 14 AA.
AC P01016;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE ANGIOTENSINOGEN (FRAGMENT).
GN AGT.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
RN [1]
RP SEQUENCE.
RA Skeggs L.T. Jr., Kahn J.R., Lentz K., Shumway N.P.;
RT "The preparation, purification, and amino acid sequence of a
RT polypeptide renin substrate.";
RL J. Exp. Med. 106:439-453(1957).
CC -!- FUNCTION: IN RESPONSE TO LOWERED BLOOD PRESSURE, THE ENZYME RENIN
CC CLEAVES ANGIOTENSIN I, FROM ANGIOTENSINOGEN. ACE (ANGIOTENSIN
CC CONVERTING ENZYME) THEN REMOVES A DIPEPTIDE TO YIELD THE
CC PHYSIOLOGICALLY ACTIVE PEPTIDE ANGIOTENSIN II, THE MOST POTENT
CC PRESSOR SUBSTANCE KNOWN, WHICH HELPS REGULATE VOLUME AND MINERAL
CC BALANCE OF BODY FLUIDS.
CC -!- TISSUE SPECIFICITY: MADE IN THE LIVER & SECRETED IN THE PLASMA.
CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY.
DR PIR; A01250; A01250.
DR INTERPRO; IPR000215; -.
DR PROSITE; PS00284; SERPIN; PARTIAL.
KW Vasoconstrictor; Plasma; Serpin.
FT PEPTIDE 1 10 ANGIOTENSIN I.
FT PEPTIDE 1 8 ANGIOTENSIN II.
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1759 MW; 2E9921F8EEEFBDD7 CRC64;

Query Match 21.6%; Score 24; DB 1; Length 14;
Best Local Similarity 83.3%; Pred. No. 4.6e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      10 PEHLLV 15
           | ||||
Db       7 PFHLLV 12
```

RESULT 2
CAQS_RAT
ID CAQS_RAT STANDARD; PRT; 20 AA.
AC P19633;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE CALSEQUESTRIN, SKELETAL MUSCLE ISOFORM (ASPARTACTIN) (LAMININ-BINDING
DE PROTEIN) (FRAGMENT).
GN CASQ1.
OS Rattus norvegicus (Rat).

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88331073.
RA   Hall D.E., Frazer K.A., Hann B.C., Reichardt L.F.;
RT   "Isolation and characterization of a laminin-binding protein from rat
RT   and chick muscle.";
RL   J. Cell Biol. 107:687-697(1988).
CC   -!- FUNCTION: CALSEQUESTRIN IS A HIGH-CAPACITY, MODERATE AFFINITY,
CC       CALCIUM-BINDING PROTEIN AND THUS ACTS AS AN INTERNAL CALCIUM STORE
CC       IN MUSCLE. THE RELEASE OF CALCIUM BOUND TO CALSEQUESTRIN THROUGH
CC       A CALCIUM RELEASE CHANNEL TRIGGERS MUSCLE CONTRACTION. BINDS 40
CC       TO 50 MOLES OF CALCIUM. ALSO BINDS LAMININ.
CC   -!- SUBCELLULAR LOCATION: THIS ISOFORM OF CALSEQUESTRIN OCCURS IN THE
CC       SARCOPLASMIC RETICULUM'S TERMINAL CISTERNAE LUMINAL SPACES OF
CC       FAST SKELETAL MUSCLE CELLS. ASPARTACTIN IS FOUND IN THE BASAL
CC       LAMINA SURROUNDING INDIVIDUAL MUSCLE FIBERS.
CC   -!- TISSUE SPECIFICITY: SKELETAL AND HEART MUSCLE.
CC   -!- SIMILARITY: BELONGS TO THE CALSEQUESTRIN FAMILY.
DR   PIR; A31049; A31049.
DR   INTERPRO; IPR001393; -.
DR   PROSITE; PS00864; CALSEQUESTRIN_2; PARTIAL.
DR   PROSITE; PS00863; CALSEQUESTRIN_1; 1.
KW   Muscle; Glycoprotein; Calcium-binding.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2238 MW;  92ADE04FC2A69280 CRC64;


  Query Match             19.8%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 FPEHLLVD 16
           |||:  ||
Db       6 FPEYDGVD 13


RESULT   3
COXN_THUOB
ID   COXN_THUOB     STANDARD;      PRT;    20 AA.
AC   P80980;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIB-HEART (EC 1.9.3.1) (FRAGMENT).
OS   Thunnus obesus (Bigeye tuna).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC   Scombridae; Thunnus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 97454291.
RA   Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA   Kadenbach B.;
RT   "The subunit structure of cytochrome-c oxidase from tuna heart and
RT   liver.";
RL   Eur. J. Biochem. 248:99-103(1997).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC       CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC       MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC       4 FERRICYTOCHROME C.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2303 MW;  0A33BD34006E5AA6 CRC64;


  Query Match             19.8%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       6 DFGFPEH 12
           | | | |
Db      14 DNGMPVH 20


RESULT   4
SAMP_MUSCA
ID   SAMP_MUSCA     STANDARD;      PRT;     9 AA.
AC   P19095;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   SERUM AMYLOID P-COMPONENT (SAP) (FRAGMENT).
OS   Mustelus canis (Smooth dogfish).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC   Mustelus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 83160932.
RA   Robey F.A., Tanaka T., Liu T.-Y.;
RT   "Isolation and characterization of two major serum proteins from the
RT   dogfish, Mustelus canis, C-reactive protein and amyloid P
RT   component.";
RL   J. Biol. Chem. 258:3889-3894(1983).
CC   -!- SUBUNIT: HOMOPENTAMER. PENTAXIN (OR PENTRAXIN) HAVE A DISCOID
CC       ARRANGEMENT OF 5 NONCOVALENTLY BOUND SUBUNITS.
CC   -!- DISEASE: SAP IS A PRECURSOR OF AMYLOID COMPONENT P WHICH IS FOUND
CC       IN BASEMENT MEMBRANE AND ASSOCIATED WITH AMYLOID DEPOSITS.
CC   -!- SIMILARITY: BELONGS TO THE PENTAXIN FAMILY.
DR   PIR; B20569; B20569.
DR   INTERPRO; IPR001759; -.
DR   PROSITE; PS00289; PENTAXIN; PARTIAL.
KW   Amyloid; Glycoprotein; Plasma; Pentaxin.
FT   DOMAIN       1      >9       PENTAXIN.
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  965 MW;  D05B5735B3386769 CRC64;


  Query Match             18.9%;  Score 21;  DB 1;  Length 9;
  Best Local Similarity   50.0%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       8 GFPEHLLV 15
           |||   |:
Db       1 GFPGKSLI 8


RESULT   5
CH60_DROME
ID   CH60_DROME     STANDARD;      PRT;    11 AA.
AC   P35380;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   MITOCHONDRIAL MATRIX PROTEIN P1 (60 KDA CHAPERONIN) (HEAT SHOCK
DE   PROTEIN 60) (HSP-60) (PROTEIN CPN60) (GROEL PROTEIN) (FRAGMENT).
GN   MMP-P1 OR HSP60.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VALLECAS; TISSUE=WING IMAGINAL DISK;
RX   MEDLINE; 93272852.
RA   Santaren J.F., van Damme J., Puype M., Vandekerckhove J.,
RA   Garcia-Bellido A.;
RT   "Identification of Drosophila wing imaginal disc proteins by two-
RT   dimensional gel analysis and microsequencing.";
RL   Exp. Cell Res. 206:220-226(1993).
CC   -!- FUNCTION: IMPLICATED IN MITOCHONDRIAL PROTEIN IMPORT AND
```

```
CC       MACROMOLECULAR ASSEMBLY. MAY FACILITATE THE CORRECT FOLDING OF
CC       IMPORTED PROTEINS. MAY ALSO PREVENT MISFOLDING AND PROMOTE THE
CC       REFOLDING AND PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED
CC       UNDER STRESS CONDITIONS IN THE MITOCHONDRIAL MATRIX (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
DR   FLYBASE; FBgn0010375; Mmp-P1.
DR   INTERPRO; IPR001844; -.
DR   PROSITE; PS00296; CHAPERONINS_CPN60; PARTIAL.
KW   Chaperone; ATP-binding; Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1243 MW;  78501A366365A6DB CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 11;
  Best Local Similarity   18.2%;  Pred. No. 1.5e+03;
  Matches    2;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPE 11
          ::::  :| |:
Db      1 VIIEQSWGSPK 11


RESULT   6
TEML_RANTE
ID   TEML_RANTE     STANDARD;      PRT;    11 AA.
AC   P56923;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN L.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES      11     11       AMIDATION.
SQ   SEQUENCE   11 AA;  1194 MW;  1E990549B3372724 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 11;
  Best Local Similarity   55.6%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     13 LLVDFLQSL 21
          || : |:||
Db      1 LLPNLLKSL 9


RESULT   7
DCMM_PSECF
ID   DCMM_PSECF     STANDARD;      PRT;    14 AA.
AC   P19914;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydoflava.
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC   Hydrogenophaga.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0142; PL0142.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1756 MW;  65583C6D1FB7C25B CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 14;
  Best Local Similarity   15.4%;  Pred. No. 2e+03;
  Matches    2;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 LLLQMDFGFPEHL 13
          :: : ::  |:|:
Db      2 MIPRFEYHAPKHV 14


RESULT   8
MDH_SYNY4
ID   MDH_SYNY4      STANDARD;      PRT;    16 AA.
AC   P80460;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Synechocystis sp. (strain PCC 6714).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN   [1]
RP   SEQUENCE.
RA   Naterstad K., Synstad B., Sirevag R.;
RL   Submitted (SEP-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   INTERPRO; IPR001252; -.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1780 MW;  61D1896F14E81984 CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   66.7%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 LQMDFG 8
          |  |||
Db      9 LTTDFG 14


RESULT   9
ATPB_PHYPA
ID   ATPB_PHYPA     STANDARD;      PRT;    21 AA.
AC   P80658;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ATP SYNTHASE BETA CHAIN (EC 3.6.1.34) (FRAGMENT).
GN   ATPB.
OS   Physcomitrella patens (Moss).
```

```
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Bryophyta; Bryopsida;
OC   Funariidae; Funariales; Funariaceae; Physcomitrella.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=PROTONEMA;
RX   MEDLINE; 97275459.
RA   Kasten B., Buck F., Nuske J., Reski R.;
RT   "Cytokinin affects nuclear- and plastome-encoded energy-converting
RT   plastid enzymes.";
RL   Planta 201:261-272(1997).
CC   -!- FUNCTION: PRODUCES ATP FROM ADP IN THE PRESENCE OF A PROTON
CC       GRADIENT ACROSS THE MEMBRANE. THE BETA CHAIN IS THE CATALYTIC
CC       SUBUNIT.
CC   -!- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC       CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC       SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC       HAS THREE MAIN SUBUNITS: A, B AND C.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE ATPASE ALPHA/BETA CHAINS FAMILY.
DR   MENDEL; 13023; PHYpa;atpB;1.
DR   INTERPRO; IPR000194; -.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; PARTIAL.
KW   ATP synthesis; Chloroplast; Thylakoid membrane; CF(1);
KW   Hydrolase; ATP-binding; Hydrogen ion transport.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2298 MW;  9558E4F5AC89D81A CRC64;


  Query Match             18.0%;  Score 20;  DB 1;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFP 10
           :|  ||
Db      16 LDIXFP 21


RESULT  10
DPP4_BOVIN
ID   DPP4_BOVIN     STANDARD;      PRT;    20 AA.
AC   P81425;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   DIPEPTIDYL PEPTIDASE IV (EC 3.4.14.5) (DPP IV) (ADENOSINE DEAMINASE
DE   COMPLEXING PROTEIN) (ADCP-I) (FRAGMENTS).
GN   DPP4.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=KIDNEY;
RX   MEDLINE; 98293306.
RA   Ben-Shooshan I., Parola A.H.;
RT   "The CP-I subunit of adenosine deaminase complexing protein from calf
RT   kidney is identical to human, mouse, and rat dipeptidyl peptidase
RT   IV.";
RL   Comp. Biochem. Physiol. 119B:289-292(1998).
CC   -!- FUNCTION: REMOVES N-TERMINAL DIPEPTIDES SEQUENTIALLY FROM
CC       POLYPEPTIDES HAVING UNSUBSTITUTED N-TERMINI PROVIDED THAT THE
CC       PENULTIMATE RESIDUE IS PROLINE. BINDS AND REGULATES THE ACTIVITY
CC       OF ADA.
CC   -!- CATALYTIC ACTIVITY: DIPEPTIDYL-POLYPEPTIDE + H(2)O = DIPEPTIDE +
CC       POLYPEPTIDE.
CC   -!- SUBUNIT: HOMODIMER.
CC   -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS IN
CC       A SOLUBLE FORM (BY SIMILARITY).
CC   -!- PTM: THE SOLUBLE FORM (SDPP) DERIVES FROM THE MEMBRANE FORM (MDPP)
CC       BY PROTEOLYTIC PROCESSING.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S9B; ALSO KNOWN AS THE
CC       PROLYL OLIGOPEPTIDASE FAMILY.
DR   INTERPRO; IPR002471; -.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW   Hydrolase; Dipeptidase; Serine protease; Transmembrane; Glycoprotein;
KW   Signal-anchor.
FT   NON_TER      1      1
FT   NON_CONS     10     11
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2356 MW;  28B78E62639F7276 CRC64;


  Query Match             17.6%;  Score 19.5;  DB 1;  Length 20;
  Best Local Similarity   33.3%;  Pred. No. 3.4e+03;
  Matches    5;  Conservative    6;  Mismatches    1;  Indels    3;  Gaps    1;

Qy       1 LLLQMDFGFPEHLLV 15
           |||::   : |:::|
Db       5 LLLEV---YTENIIV 16


RESULT  11
OXLA_OPHHA
ID   OXLA_OPHHA     STANDARD;      PRT;    19 AA.
AC   P81383;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   L-AMINO ACID OXIDASE (EC 1.4.3.2) (LAO) (LAAO) (FRAGMENT).
OS   Ophiophagus hannah (King cobra) (Naja hannah).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Elapidae; Elapinae; Ophiophagus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 94361525.
RA   Ponnudurai G., Chung M.C.M., Tan N.-H.;
RT   "Purification and properties of the L-amino acid oxidase from Malayan
RT   pit viper (Calloselasma rhodostoma) venom.";
RL   Arch. Biochem. Biophys. 313:373-378(1994).
RN   [2]
RP   SEQUENCE OF 1-15.
RC   TISSUE=VENOM;
RX   MEDLINE; 97449790.
RA   Ahn M.Y., Lee B.M., Kim Y.S.;
RT   "Characterization and cytotoxicity of L-amino acid oxidase from the
RT   venom of king cobra (Ophiophagus hannah).";
RL   Int. J. Biochem. Cell Biol. 29:911-919(1997).
CC   -!- FUNCTION: HAS CYTOTOXIC ACTIVITY (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: A L-AMINO ACID + H(2)O + O(2) = A 2-OXO ACID +
CC       NH(3) + H(2)O(2).
CC   -!- COFACTOR: FAD.
CC   -!- SUBUNIT: HOMODIMER (PROBABLE).
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE FLAVIN MONOAMINE OXIDASE FAMILY.
CC       STRONG, TO MOUSE FIG-1.
KW   Oxidoreductase; Flavoprotein; FAD; Glycoprotein; Venom.
FT   CONFLICT     1      1       H -> S (IN REF. 2).
FT   NON_TER     19     19
SQ   SEQUENCE   19 AA;  2298 MW;  DD911A5B414F1427 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 3.9e+03;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LQMDFGFPEH 12
           |:  |  ||:
Db       5 LEESFQEPEY 14


RESULT  12
UP21_UPEIN
```

```
ID   UP21_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82027;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.1.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST
CC       L.MESENTERIODES, M.LUTEUS AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1926; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1927 MW;  328834D77BA353D2 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKKV 9


RESULT  13
UP25_UPEIN
ID   UP25_UPEIN     STANDARD;      PRT;    19 AA.
AC   P82031;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   UPERIN 2.5.
OS   Uperoleia inundata (Floodplain toadlet).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC   Uperoleia.
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Bradford A.M., Raftery M.J., Bowie J.H., Tyler M.J., Wallace J.C.,
RA   Adams G.W., Severini C.;
RT   "Novel uperin peptides from the dorsal glands of the australian
RT   floodplain toadlet Uperoleia inundata.";
RL   Aust. J. Chem. 49:475-484(1996).
CC   -!- FUNCTION: SHOWS A MEDIUM ANTIBACTERIAL ACTIVITY AGAINST M.LUTEUS,
CC       L.MESENTERIODES AND S.UBERIS.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=1940; METHOD=FAB.
KW   Amphibian skin; Antibiotic.
SQ   SEQUENCE   19 AA;  1941 MW;  5E94C6C757B463D9 CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 3.9e+03;
  Matches    3;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 LVDFLQSL 21
           :||| : :
Db       2 IVDFAKGV 9


RESULT  14
SUCB_CANFA
ID   SUCB_CANFA     STANDARD;      PRT;    20 AA.
AC   P99507;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   PUTATIVE SUCCINYL-COA LIGASE [GDP-FORMING] BETA-CHAIN (EC 6.2.1.4)
DE   (SUCCINYL-COA SYNTHETASE, BETA CHAIN) (SCS-BETA) (FRAGMENT).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=HEART;
RX   MEDLINE; 98163340.
RA   Dunn M.J., Corbett J.M., Wheeler C.H.;
RT   "HSC-2DPAGE and the two-dimensional gel electrophoresis database of
RT   dog heart proteins.";
RL   Electrophoresis 18:2795-2802(1997).
CC   -!- CATALYTIC ACTIVITY: SUCCINATE + COA + GTP = SUCCINYL-COA + GDP +
CC       ORTHOPHOSPHATE.
CC   -!- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC
CC       ACID CYCLE.
CC   -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL (BY SIMILARITY).
CC   -!- SIMILARITY: TO OTHER BETA SUBUNITS OF SUCCINYL-COA SYNTHETASE,
CC       OF MALATE--COA LIGASE AND TO ATP CITRATE-LYASE.
DR   HSC-2DPAGE; P99507; DOG.
DR   INTERPRO; IPR000303; -.
DR   PROSITE; PS01217; SUCCINYL_COA_LIG_3; PARTIAL.
KW   Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2248 MW;  BE8AEFD54DBDAC2E CRC64;


  Query Match             17.1%;  Score 19;  DB 1;  Length 20;
  Best Local Similarity   33.3%;  Pred. No. 4.1e+03;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      11 EHLLVDFLQ 19
           |:: :: ||
Db       5 EYMSMELLQ 13


RESULT  15
DCMS_PSECA
ID   DCMS_PSECA     STANDARD;      PRT;    21 AA.
AC   P19921;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] SMALL CHAIN (EC 1.2.2.4)
DE   (FRAGMENT).
OS   Pseudomonas carboxydovorans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Oligotropha.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=OM5;
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 = CO(2) +
CC       2 H(+) + FERRICYTOCHROME B-561.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0144; PL0144.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2270 MW;  68D4380629401B9C CRC64;
```

Query Match 17.1%; Score 19; DB 1; Length 21;
Best Local Similarity 44.4%; Pred. No. 4.3e+03;
Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```
Qy       8 GFPEHLLVD 16
           | |   ||:
Db      12 GHPVEALVE 20
```

Search completed: December 21, 2000, 08:38:16
Job time: 149 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:37:55 ; Search time 37.32 Seconds
(without alignments)
55.044 Million cell updates/sec

Title: US-08-934-367-34
Perfect score: 111
Sequence: 1 LLLQMDFGFPEHLLVDFLQSLS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4962

Minimum DB seq length: 0
Maximum DB seq length: 22

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	27.9	19	6	Q9TRR6	Q9trr6 oryctolagus
2	30	27.0	21	2	P70861	P70861 borrelia bu
3	29	26.1	20	5	Q9TWH5	Q9twh5 heliothis v
4	27	24.3	12	4	Q9UGS1	Q9ugs1 homo sapien
5	27	24.3	14	6	Q9TRQ7	Q9trq7 bos taurus
6	27	24.3	22	10	Q9S8E1	Q9s8e1 capsicum an
7	26.5	23.9	21	2	Q9ZG55	Q9zg55 chlamydia t
8	26	23.4	22	6	O02830	O02830 oryctolagus
9	25	22.5	16	3	O94554	O94554 schizosacch
10	25	22.5	19	13	Q9PS70	Q9ps70 gallus gall
11	25	22.5	20	10	Q9S878	Q9s878 petunia hyb
12	24.5	22.1	22	13	Q9PS42	Q9ps42 gallus gall
13	24	21.6	13	2	O31364	O31364 borrelia ga
14	24	21.6	13	2	O31365	O31365 borrelia ga
15	24	21.6	13	2	O34622	O34622 borrelia bu
16	24	21.6	18	1	Q9UYK7	Q9uyk7 pyrococcus
17	24	21.6	19	5	Q9W5Q8	Q9w5q8 drosophila
18	24	21.6	20	6	Q9TTG3	Q9ttg3 ateles belz
19	24	21.6	21	2	Q9X3D0	Q9x3d0 prochloroco
20	23	20.7	8	2	O85406	O85406 coxiella bu
21	23	20.7	9	5	O96417	O96417 drosophila
22	23	20.7	10	2	Q9X534	Q9x534 leclercia a
23	23	20.7	10	2	Q9X533	Q9x533 escherichia
24	23	20.7	13	2	O31296	O31296 borrelia af
25	23	20.7	13	2	O34770	O34770 borrelia af
26	23	20.7	13	11	O35758	O35758 rattus norv
27	23	20.7	17	2	Q9X3I3	Q9x3i3 prochloroco
28	23	20.7	18	13	P82068	P82068 litoria gen
29	23	20.7	19	11	Q9QVE3	Q9qve3 rattus sp.
30	23	20.7	20	13	Q9PRN5	Q9prn5 oncorhynchu
31	23	20.7	21	4	Q16575	Q16575 homo sapien
32	23	20.7	21	8	Q9ZYB7	Q9zyb7 spinaria sp
33	23	20.7	21	10	Q41496	Q41496 solanum tub
34	23	20.7	21	12	Q80817	Q80817 human t-cel
35	23	20.7	22	1	P71537	P71537 methanosarc
36	22.5	20.3	17	6	Q9TR11	Q9tr11 bos taurus
37	22	19.8	13	4	Q9UPE7	Q9upe7 homo sapien
38	22	19.8	14	2	Q56945	Q56945 yersinia ps
39	22	19.8	15	10	Q41526	Q41526 triticum ae
40	22	19.8	17	8	Q9T2S2	Q9t2s2 solanum tub
41	22	19.8	17	11	Q9QUJ4	Q9quj4 mus sp. mep
42	22	19.8	19	9	Q38371	Q38371 bacteriopha
43	22	19.8	20	5	Q9TWI4	Q9twi4 ceratitis c
44	22	19.8	20	6	Q9TRA1	Q9tra1 bos taurus
45	22	19.8	20	13	Q9PRM4	Q9prm4 gallus gall

ALIGNMENTS

RESULT 1
Q9TRR6
ID Q9TRR6 PRELIMINARY; PRT; 19 AA.
AC Q9TRR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CALCYCLIN-ASSOCIATED PROTEIN, CAP50=CA2+/PHOSPHOLIPID-BINDING PROTEIN
DE L-14 FRAGMENT.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 92250478.
RA Tokumitsu H., Mizutani A., Minami H., Kobayashi R., Hidaka H.;
RT "A calcyclin-associated protein is a newly identified member of the
RT Ca2+/phospholipid-binding proteins, annexin family.";
RL J. Biol. Chem. 267:8919-8924(1992).
SQ SEQUENCE 19 AA; 2018 MW; 9A54062504B8322E CRC64;

Query Match 27.9%; Score 31; DB 6; Length 19;
Best Local Similarity 43.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

```
Qy       7 FGFPEHLLVDFLQSLS 22
           ||  |  ::| | | |
Db       2 FGTDEQAIIDXLGSRS 17
```

RESULT 2
P70861
ID P70861 PRELIMINARY; PRT; 21 AA.
AC P70861;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE THDF (FRAGMENT).
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]

```
RP   SEQUENCE FROM N.A.
RC   STRAIN=212;
RX   MEDLINE; 97312006.
RA   Ge Y., Old I.G., Girons I.S., Charon N.W.;
RT   "The flgK motility operon of Borrelia burgdorferi is initiated by a
RT   sigma 70-like promoter.";
RL   Microbiology 143:1681-1690(1997).
DR   EMBL; U62901; AAB62742.1; -.
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2432 MW;  F33E1EC548BD5B33 CRC64;


  Query Match             27.0%;  Score 30;  DB 2;  Length 21;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+02;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 MDFGFPEHLL 14
           ::| ||| :|
Db       7 IEFDFPEGIL 16


RESULT   3
Q9TWH5
ID   Q9TWH5      PRELIMINARY;      PRT;    20 AA.
AC   Q9TWH5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   40 KDA GAP JUNCTION PROTEIN (FRAGMENT).
OS   Heliothis virescens (Noctuid moth) (Owlet moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Noctuoidea; Noctuidae; Heliothinae; Heliothis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95347000.
RA   Ryerse J.S.;
RT   "Immunocytochemical, electrophoresis, and immunoblot analysis of
RT   Heliothis virescens gap junctions isolated in the presence and absence
RT   of protease inhibitors.";
RL   Cell Tissue Res. 281:179-186(1995).
SQ   SEQUENCE   20 AA;  2304 MW;  A298D3EB3E89586B CRC64;


  Query Match             26.1%;  Score 29;  DB 5;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 3.5e+02;
  Matches    6;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGFPEHLLVDF 17
           ::  :| |: | |  ||
Db       3 VIFNIDGGYLEFLTRDF 19


RESULT   4
Q9UGS1
ID   Q9UGS1      PRELIMINARY;      PRT;    12 AA.
AC   Q9UGS1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ796I17.4 (NOVEL PROTEIN SIMILAR TO GS2) (FRAGMENT).
GN   DJ796I17.4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Barlow K.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL035398; CAB63074.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   12 AA;  1430 MW;  AF7740ABECB69AA6 CRC64;


  Query Match             24.3%;  Score 27;  DB 4;  Length 12;
  Best Local Similarity   44.4%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLQMDFGF 9
           ::|:||  |
Db       1 MILEMDMSF 9


RESULT   5
Q9TRQ7
ID   Q9TRQ7      PRELIMINARY;      PRT;    14 AA.
AC   Q9TRQ7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CALCYCLIN-ASSOCIATED PROTEIN PEPTIDE L-8, CAP-50=ANNEXIN.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92317074.
RA   Mizutani A., Usuda N., Tokumitsu H., Minami H., Yasui K.,
RA   Kobayashi R., Hidaka H.;
RT   "CAP-50, a newly identified annexin, localizes in nuclei of cultured
RT   fibroblast 3Y1 cells.";
RL   J. Biol. Chem. 267:13498-13504(1992).
SQ   SEQUENCE   14 AA;  1446 MW;  C8322EB96DD9C6C6 CRC64;


  Query Match             24.3%;  Score 27;  DB 6;  Length 14;
  Best Local Similarity   41.7%;  Pred. No. 4.9e+02;
  Matches    5;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFL 18
           ||  |  ::| |
Db       2 FGTDEQAIIDXL 13


RESULT   6
Q9S8E1
ID   Q9S8E1      PRELIMINARY;      PRT;    22 AA.
AC   Q9S8E1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ANNEXIN (FRAGMENT).
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Capsicum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95353219.
RA   Hoshino T., Mizutani A., Chida M., Hidaka H., Mizutani J.;
RT   "Plant annexin form homodimer during Ca(2+)-dependent liposome
RT   aggregation.";
RL   Biochem. Mol. Biol. Int. 35:749-755(1995).
SQ   SEQUENCE   22 AA;  2465 MW;  A36D11A1FE311F0F CRC64;


  Query Match             24.3%;  Score 27;  DB 10;  Length 22;
  Best Local Similarity   41.7%;  Pred. No. 8e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       7 FGFPEHLLVDFL 18
           :|  | |::| |
Db       3 WGTDEKLIIDIL 14
```

```
RESULT   7
Q9ZG55
ID   Q9ZG55      PRELIMINARY;      PRT;    21 AA.
AC   Q9ZG55;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   ATP-BINDING PROTEIN (FRAGMENT).
GN   RECF.
OS   Chlamydia trachomatis.
OC   Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=L2 434B;
RA   Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT   "Gene identification of Chlamydia trachomatis by random DNA
RT   sequencing.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF087306; AAD04082.1; -.
KW   ATP-binding.
FT   NON_TER       1      1
FT   NON_TER      21     21
SQ   SEQUENCE   21 AA;  2336 MW;  0185D9AC428276D9 CRC64;


  Query Match             23.9%;  Score 26.5;  DB 2;  Length 21;
  Best Local Similarity   41.7%;  Pred. No. 9.2e+02;
  Matches    5;  Conservative    4;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      2 LLQMDFGFPEHL 13
          :||:   ||:|:
Db      8 ILQL-ISFPKHI 18


RESULT   8
O02830
ID   O02830      PRELIMINARY;      PRT;    22 AA.
AC   O02830;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PRO ALPHA1 TYPE II COLLAGEN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 96377339.
RA   Metsaranta M., Kujala U.M., Pelliniemi L., Osterman H., Aho H.,
RA   Vuorio E.;
RT   "Evidence for insufficient chondrocytic differentiation during repair
RT   of full-thickness defects of articular cartilage.";
RL   Matrix Biol. 15:39-47(1996).
DR   EMBL; S83370; AAB50773.1; -.
DR   INTERPRO; IPR000885; -.
DR   PFAM; PF01410; COLFI; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   22 AA;  2292 MW;  2E05FE169844236D CRC64;


  Query Match             23.4%;  Score 26;  DB 6;  Length 22;
  Best Local Similarity   71.4%;  Pred. No. 1.2e+03;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 MDFGFPE 11
          || | ||
Db      3 MDIGGPE 9


RESULT   9
O94554
ID   O94554      PRELIMINARY;      PRT;    16 AA.
AC   O94554;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   CALMODULIN KINASE 2 (FRAGMENT).
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
OC   Schizosaccharomycetaceae; Schizosaccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Alemany V., Aligue R.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U57982; AAD09466.1; -.
KW   Kinase.
FT   NON_TER       1      1
SQ   SEQUENCE   16 AA;  1846 MW;  4A673B1F4C3288D9 CRC64;


  Query Match             22.5%;  Score 25;  DB 3;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 FGFP 10
          ||||
Db      5 FGFP 8


RESULT  10
Q9PS70
ID   Q9PS70      PRELIMINARY;      PRT;    19 AA.
AC   Q9PS70;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   LOW DENSITY LIPOPROTEIN RECEPTOR-RELATED PROTEIN.
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92011685.
RA   Stifani S., Barber D.L., Aebersold R., Steyrer E., Shen X., Nimpf J.,
RA   Schneider W.J.;
RT   "The laying hen expresses two different low density lipoprotein
RT   receptor-related proteins.";
RL   J. Biol. Chem. 266:19079-19087(1991).
SQ   SEQUENCE   19 AA;  1861 MW;  4EEC931205620608 CRC64;


  Query Match             22.5%;  Score 25;  DB 13;  Length 19;
  Best Local Similarity   40.0%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    1;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      2 LLQMDFGFPEHLLVD 16
          ||    | |  | :|
Db      1 LLAQGLGXPTALALD 15


RESULT  11
Q9S878
ID   Q9S878      PRELIMINARY;      PRT;    20 AA.
AC   Q9S878;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   NADPH-CYTOCHROME P450 REDUCTASE (EC 1.6.2.4) (FRAGMENT).
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Petunia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95083756.
```

```
RA   Menting J.G., Cornish E., Scopes R.K.;
RT   "Purification and partial characterization of NADPH-cytochrome c
RT   reductase from Petunia hybrida flowers.";
RL   Plant Physiol. 106:643-650(1994).
SQ   SEQUENCE   20 AA;  2390 MW;  8BCDA6F8CF7EEDEE CRC64;


  Query Match             22.5%;  Score 25;  DB 10;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 QMDFGFPEHL 13
           :||| : | |
Db       1 RMDFIYEEEL 10


RESULT  12
Q9PS42
ID   Q9PS42      PRELIMINARY;      PRT;    22 AA.
AC   Q9PS42;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PHENOBARBITAL-INDUCED 48 KDA CYTOCHROME P-450 (FRAGMENT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92406903.
RA   Nakai K., Ward A.M., Gannon M., Rifkind A.B.;
RT   "Beta-naphthoflavone induction of a cytochrome P-450 arachidonic acid
RT   epoxygenase in chick embryo liver distinct from the aryl hydrocarbon
RT   hydroxylase and from phenobarbital-induced arachidonate epoxygenase.";
RL   J. Biol. Chem. 267:19503-19512(1992).
SQ   SEQUENCE   22 AA;  2377 MW;  6BA9C990EE33E9C1 CRC64;


  Query Match             22.1%;  Score 24.5;  DB 13;  Length 22;
  Best Local Similarity   58.3%;  Pred. No. 2e+03;
  Matches    7;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy       5 MDF-GFPEHLLV 15
           ||| | |  ||:
Db       1 MDFLGLPTILLL 12


RESULT  13
O31364
ID   O31364      PRELIMINARY;      PRT;    13 AA.
AC   O31364;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=G25;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93700; AAC45534.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1449 MW;  08FAB9BB0E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


RESULT  14
O31365
ID   O31365      PRELIMINARY;      PRT;    13 AA.
AC   O31365;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia garinii.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IP90;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
DR   EMBL; U93701; AAC45536.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1382 MW;  08EFB4180E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


RESULT  15
O34622
ID   O34622      PRELIMINARY;      PRT;    13 AA.
AC   O34622;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   GUAA (FRAGMENT).
GN   GUAA.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT39 (ILL-1), DN127CL9-2, AND 25015;
RX   MEDLINE; 97426044.
RA   Tilly K., Casjens S., Stevenson B., Bono J.L., Samuels D.S., Hogan D.,
RA   Rosa P.;
RT   "The Borrelia burgdorferi circular plasmid cp26: conservation of
RT   plasmid structure and targeted inactivation of the ospC gene.";
RL   Mol. Microbiol. 25:361-374(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DN127CL9-2, AND 25015;
RX   MEDLINE; 95154673.
RA   Stevenson B., Barthold S.W.;
RT   "Expression and sequence of outer surface protein C among North
RT   American isolates of Borrelia burgdorferi.";
RL   FEMS Microbiol. Lett. 124:367-372(1994).
DR   EMBL; U93694; AAC45522.1; -.
DR   EMBL; U04280; AAC45543.1; -.
```

```
DR   EMBL; U04282; AAC45541.1; -.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1407 MW;  08FAB3930E0BC720 CRC64;


  Query Match             21.6%;  Score 24;  DB 2;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 LLQMDFG 8
           :| :|||
Db       6 ILVLDFG 12


Search completed: December 21, 2000, 08:37:55
Job time: 288 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:50 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-35
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	34.4	9	1	US-08-454-207A-43	Sequence 43, Appl
2	32	33.3	17	5	5266328-11	Patent No. 5266328
3	29	30.2	18	3	US-09-100-414B-26	Sequence 26, Appl
4	29	30.2	18	3	US-09-100-414B-30	Sequence 30, Appl
5	28	29.2	19	1	US-07-654-839-5	Sequence 5, Appli
6	27	28.1	9	1	US-08-454-207A-45	Sequence 45, Appl
7	26	27.1	9	1	US-08-454-207A-30	Sequence 30, Appl
8	26	27.1	9	1	US-08-454-207A-41	Sequence 41, Appl
9	26	27.1	9	1	US-08-454-207A-44	Sequence 44, Appl
10	26	27.1	14	4	PCT-US95-06077-5	Sequence 5, Appli
11	26	27.1	15	2	US-08-195-874-2	Sequence 2, Appli
12	26	27.1	15	2	US-08-484-905-26	Sequence 26, Appl
13	26	27.1	15	3	US-08-481-985B-26	Sequence 26, Appl
14	26	27.1	15	4	PCT-US95-01671-2	Sequence 2, Appli
15	26	27.1	16	1	US-08-305-871A-9	Sequence 9, Appli
16	26	27.1	20	1	US-08-218-025A-81	Sequence 81, Appl
17	25	26.0	9	1	US-08-454-207A-29	Sequence 29, Appl
18	25	26.0	9	1	US-08-454-207A-31	Sequence 31, Appl
19	25	26.0	9	1	US-08-454-207A-32	Sequence 32, Appl
20	25	26.0	9	1	US-08-454-207A-33	Sequence 33, Appl
21	25	26.0	9	1	US-08-454-207A-34	Sequence 34, Appl
22	25	26.0	9	1	US-08-454-207A-35	Sequence 35, Appl
23	25	26.0	9	1	US-08-454-207A-36	Sequence 36, Appl
24	25	26.0	9	1	US-08-454-207A-37	Sequence 37, Appl
25	25	26.0	9	1	US-08-454-207A-38	Sequence 38, Appl
26	25	26.0	9	1	US-08-454-207A-39	Sequence 39, Appl
27	25	26.0	9	1	US-08-454-207A-40	Sequence 40, Appl
28	25	26.0	9	3	US-08-159-339A-78	Sequence 78, Appl
29	25	26.0	10	1	US-08-454-207A-2	Sequence 2, Appli
30	25	26.0	10	1	US-08-454-207A-5	Sequence 5, Appli
31	25	26.0	10	2	US-08-456-112B-42	Sequence 42, Appl
32	25	26.0	10	3	US-08-159-339A-77	Sequence 77, Appl
33	25	26.0	11	2	US-08-310-912A-117	Sequence 117, App
34	25	26.0	11	4	PCT-US95-04589-117	Sequence 117, App
35	25	26.0	13	4	PCT-US94-01234-47	Sequence 47, Appl
36	25	26.0	14	2	US-09-133-774-9	Sequence 9, Appli
37	25	26.0	14	3	US-09-303-862-9	Sequence 9, Appli
38	25	26.0	17	1	US-07-992-288-4	Sequence 4, Appli
39	25	26.0	17	1	US-07-989-764-4	Sequence 4, Appli
40	25	26.0	18	1	US-08-233-203-10	Sequence 10, Appl
41	25	26.0	18	1	US-08-182-483A-25	Sequence 25, Appl
42	25	26.0	18	1	US-08-243-879A-24	Sequence 24, Appl
43	25	26.0	18	1	US-08-499-523-45	Sequence 45, Appl
44	25	26.0	18	3	US-09-100-414B-24	Sequence 24, Appl
45	25	26.0	20	2	US-08-564-972-20	Sequence 20, Appl

ALIGNMENTS

```
RESULT 1
US-08-454-207A-43
; Sequence 43, Application US/08454207A
; Patent No. 5710123
; GENERAL INFORMATION:
;   APPLICANT: Heavner, George A.
;   APPLICANT: Kruszynski, Marian
;   TITLE OF INVENTION: PEPTIDE INHIBITORS OF SELECTIN BINDING
;   NUMBER OF SEQUENCES: 70
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;     STREET: One Liberty Place - 46th Floor
;     CITY: Philadelphia
;     STATE: PA
;     COUNTRY: U.S.A.
;     ZIP: 19103
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: 3.5 inch disk, 1.44 Mb
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: WordPerfect 6.1
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/454,207A
;     FILING DATE: 09-JUN-1995
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: PCT/US93/12110
;     FILING DATE: 13-DEC-1993
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 07/997,771
;     FILING DATE: 18-DEC-1992
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Dianne B. Elderkin
;     REGISTRATION NUMBER: 28,598
;     REFERENCE/DOCKET NUMBER: CCOR-0183
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 215-568-3100
;     TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 43:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 9 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
US-08-454-207A-43


  Query Match             34.4%;  Score 33;  DB 1;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 1.2e+05;
  Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 SKKKLFL 14
```

```
          |||||||
Db      1 SKKKLFL 7


RESULT   2
5266328-11
;Patent No. 5266328
;    APPLICANT: SKUBITZ, AMY P.N.;FURCHT, LEO T.
;    TITLE OF INVENTION: LAMININ  CHAIN POLYPEPTIDES FROM
;THE CARBOXY TERMINAL GLOBULAR DOMAIN
;    NUMBER OF SEQUENCES: 14
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/573,672
;      FILING DATE: 27-AUG-1990
;SEQ ID NO:11:
;      LENGTH: 17
5266328-11


  Query Match             33.3%;  Score 32;  DB 5;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    6;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TVQASYSKKKLF 13
          ||:  | |:| |
Db      6 TVKTEYIKRKAF 17


RESULT   3
US-09-100-414B-26
; Sequence 26, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-26


  Query Match             30.2%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 KKKLFL 14
          ||||||
Db      1 KKKLFL 6


RESULT   4
US-09-100-414B-30
; Sequence 30, Application US/09100414B
; Patent No. 6025468
;  GENERAL INFORMATION:
;    APPLICANT:  Wang, Chang Yi
;    TITLE OF INVENTION:  NOVEL LHRH PEPTIDE
;    TITLE OF INVENTION:  IMMUNOGENS
;    NUMBER OF SEQUENCES:  106
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10154-0054
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC Windows
;      SOFTWARE:  Word 97
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/100,414B
;      FILING DATE:  20-JUNE-1998
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Maria H. Lin
;      REGISTRATION NUMBER:  29,323
;      REFERENCE/DOCKET NUMBER:  1151-4157
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-758-4800
;      TELEFAX:  212-751-6849
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  18 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-09-100-414B-30


  Query Match             30.2%;  Score 29;  DB 3;  Length 18;
  Best Local Similarity   100.0%;  Pred. No. 83;
  Matches    6;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      9 KKKLFL 14
          ||||||
Db      1 KKKLFL 6


RESULT   5
US-07-654-839-5
; Sequence 5, Application US/07654839
; Patent No. 5372933
;  GENERAL INFORMATION:
;    APPLICANT:  Zamarron, Concepcion
;    APPLICANT:  Plow, Edward F
;    APPLICANT:  Ginsberg, Mark H
;    TITLE OF INVENTION:  MONOCLONAL ANTIBODIES AGAINST
;    TITLE OF INVENTION:  RECEPTOR-INDUCED BINDING SITES
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  The Scripps Research Institute, Office of
;      ADDRESSEE:  Patent Counsel
;      STREET:  10666 No. 5372933th Torrey Pines Road., Suite 220,
;      STREET:  Mail Drop TPC8
;      CITY:  La Jolla
;      STATE:  CA
```

```
;       COUNTRY:  USA
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/654,839
;       FILING DATE:  19910213
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/252,753
;       FILING DATE:  03-OCT-1988
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/415,029
;       FILING DATE:  29-SEP-1989
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Bingham, Douglas A
;       REGISTRATION NUMBER:  32,457
;       REFERENCE/DOCKET NUMBER:  SCR0367P
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-554-2937
;       TELEFAX:  619-554-6312
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  19 amino acids
;       TYPE:  AMINO ACID
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
;     HYPOTHETICAL:  NO
US-07-654-839-5


  Query Match             29.2%;  Score 28;  DB 1;  Length 19;
  Best Local Similarity   38.5%;  Pred. No. 1.3e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      7 YSKKKLFLSLLDF 19
          || ||  : :: |
Db      1 YSMKKTTMKIIPF 13


RESULT   6
US-08-454-207A-45
; Sequence 45, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  45:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-45


  Query Match             28.1%;  Score 27;  DB 1;  Length 9;
  Best Local Similarity   85.7%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 SKKKLFL 14
          ||||| |
Db      1 SKKKLIL 7


RESULT   7
US-08-454-207A-30
; Sequence 30, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-30


  Query Match             27.1%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+05;
```

```
  Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLSL 16
          ||||| | :
Db      1 SKKKLALCI 9


RESULT   8
US-08-454-207A-41
; Sequence 41, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  41:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-41


  Query Match             27.1%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLS 15
          ||||| |:
Db      1 SKKKLALT 8


RESULT   9
US-08-454-207A-44
; Sequence 44, Application US/08454207A
; Patent No. 5710123
;  GENERAL INFORMATION:
;    APPLICANT:  Heavner, George A.
;    APPLICANT:  Kruszynski, Marian
;    TITLE OF INVENTION:  PEPTIDE INHIBITORS OF SELECTIN BINDING
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5710123ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/454,207A
;      FILING DATE:  09-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12110
;      FILING DATE:  13-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/997,771
;      FILING DATE:  18-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dianne B. Elderkin
;      REGISTRATION NUMBER:  28,598
;      REFERENCE/DOCKET NUMBER:  CCOR-0183
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  44:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-454-207A-44


  Query Match             27.1%;  Score 26;  DB 1;  Length 9;
  Best Local Similarity   85.7%;  Pred. No. 1.2e+05;
  Matches    6;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 SKKKLFL 14
          ||||| |
Db      1 SKKKLHL 7


RESULT  10
PCT-US95-06077-5
; Sequence 5, Application PC/TUS9506077
;  GENERAL INFORMATION:
;    APPLICANT:  Immunobiology Research, Institute Inc.
;    TITLE OF INVENTION:  Vaccine Interdiction of Extracellular
;    TITLE OF INVENTION:  Transactivating Proteins of Human Immunodeficiency Virus
;    TITLE OF INVENTION:  and Other Chronically Infecting Viruses Employing Similar
;    TITLE OF INVENTION:  Intercellular Transactivating Strategies
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Howson and Howson
;      STREET:  Spring House Corporate Cntr, PO Box 457
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06077
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/247,991
;      FILING DATE:  23-MAY-1994
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  IRI44PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-540-9200
;      TELEFAX:  215-540-5818
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
PCT-US95-06077-5


  Query Match             27.1%;  Score 26;  DB 4;  Length 14;
  Best Local Similarity   83.3%;  Pred. No. 2e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 SYSKKK 11
             || |||
Db       1 SYGKKK 6


RESULT  11
US-08-195-874-2
; Sequence 2, Application US/08195874
; Patent No. 5817308
;  GENERAL INFORMATION:
;    APPLICANT:  Scott, David W.
;    APPLICANT:  Zambidis, Elias T.
;    TITLE OF INVENTION:  Tolerogenic Fusion Proteins of
;    TITLE OF INVENTION:  Immunoglobulins and Methods for Inducing and Maintaining
;    TITLE OF INVENTION:  Tolerance
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schwegman, Lundberg & Woessner, P.A.
;      STREET:  3500 IDS Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/195,874
;      FILING DATE:  02-FEB-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mueting, Ann M.
;      REGISTRATION NUMBER:  33,977
;      REFERENCE/DOCKET NUMBER:  850.14-US-01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-339-0331
;      TELEFAX:  612-339-3061
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-195-874-2


  Query Match             27.1%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
             ::| | |||
Db       7 LKAIYEKKK 15


RESULT  12
US-08-484-905-26
; Sequence 26, Application US/08484905
; Patent No. 5976551
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
;    APPLICANT:  Kourilsky, Philippe
;    TITLE OF INVENTION:  An Altered Major Histocompatibility
;    TITLE OF INVENTION:  Complex(MHC) Determinant and Methods for Using the
;    TITLE OF INVENTION:  Determinant
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS-/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/484,905
;      FILING DATE:  07-JUNE-1995
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Potter, Jane E. R.
;      REGISTRATION NUMBER:  33,332
;      REFERENCE/DOCKET NUMBER:  03495.0106-03000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-484-905-26


  Query Match             27.1%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
             ::| | |||
Db       7 LKAIYEKKK 15


RESULT  13
US-08-481-985B-26
; Sequence 26, Application US/08481985B
; Patent No. 6011146
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
```

```
;    APPLICANT:  Kourilsky, Phillipe
;    TITLE OF INVENTION:  Altered Major Histocompatibility Complex
;    TITLE OF INVENTION:
;    NUMBER OF SEQUENCES:  148
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,985B
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/801,818
;      FILING DATE:  05-DEC-1991
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/792,473
;      FILING DATE:  15-NOV-1991
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Meyers, Kenneth J.
;      REGISTRATION NUMBER:  25,146
;      REFERENCE/DOCKET NUMBER:  03495.0106-04000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-481-985B-26


  Query Match             27.1%;  Score 26;  DB 3;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           ::| | |||
Db       7 LKAIYEKKK 15


RESULT  14
PCT-US95-01671-2
; Sequence 2, Application PC/TUS9501671
;  GENERAL INFORMATION:
;    APPLICANT:  University of Rochester
;    TITLE OF INVENTION:  Inducing Tolerance With Tolerogenic
;    TITLE OF INVENTION:  Fusion Proteins
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Schwegman, Lundberg & Woessner
;      STREET:  3500 IDS Center
;      CITY:  Minneapolis
;      STATE:  Minnesota
;      COUNTRY:  United States of America
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/01671
;      FILING DATE:  10-FEB-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kevin W. Raasch
;      REGISTRATION NUMBER:  35,651
;      REFERENCE/DOCKET NUMBER:  850.114WO1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-339-0331
;      TELEFAX:  612-339-3061
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  15 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
PCT-US95-01671-2


  Query Match             27.1%;  Score 26;  DB 4;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+02;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           ::| | |||
Db       7 LKAIYEKKK 15


RESULT  15
US-08-305-871A-9
; Sequence 9, Application US/08305871A
; Patent No. 5736142
;  GENERAL INFORMATION:
;    APPLICANT:  Sette, Alessandro
;    APPLICANT:  Gaeta, Federico
;    APPLICANT:  Grey, Howard M.
;    APPLICANT:  Sidney, John
;    APPLICANT:  Alexander, Jeffrey L.
;    TITLE OF INVENTION:  Alteration of Immune Response Using Pan
;    TITLE OF INVENTION:  DR-Binding Peptides
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend and Crew LLP
;      STREET:  Two Embarcadero Center, Eighth Floor
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94111-3834
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/305,871A
;      FILING DATE:  14-SEP-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/121,101
;      FILING DATE:  14-SEP-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bastian, Kevin L.
;      REGISTRATION NUMBER:  34,774
;      REFERENCE/DOCKET NUMBER:  14137-0062-10
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 576-0200
;      TELEFAX:  (415) 576-0300
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
```

; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-305-871A-9

Query Match 27.1%; Score 26; DB 1; Length 16;
Best Local Similarity 55.6%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
Qy      3 VQASYSKKK 11
          ::| | |||
Db      8 LKAIYEKKK 16
```

Search completed: December 21, 2000, 08:31:50
Job time: 373 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:08 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-35
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	31.2	18	2	S14560	probable heme-bind
2	28	29.2	14	2	S13864	methyl coenzyme M
3	26	27.1	14	4	I52618	hemoglobin beta ch
4	25	26.0	6	2	A61411	ameletin - rat
5	24	25.0	17	2	I49593	cystic fibrosis tr
6	24	25.0	17	2	I84733	gene CFTR protein
7	23	24.0	19	2	A49192	transthyretin - bu
8	23	24.0	20	2	S39049	cytotoxin-binding
9	22	22.9	13	2	PH0786	T-cell receptor al
10	22	22.9	15	2	S65717	prostaglandin D-sy
11	22	22.9	15	2	PA0046	protein QA100044 -
12	22	22.9	16	2	S34444	blaZ protein - Sta
13	22	22.9	18	2	S54272	CTC 75 protein - h
14	21	21.9	13	2	G22565	R-phycoerythrin ga
15	21	21.9	14	2	S29209	avenin alpha-2 - o
16	21	21.9	14	2	A44920	2-halobenzoate 1,2
17	21	21.9	15	2	PN0662	dystrophin-associa
18	21	21.9	16	2	B45895	T-cell surface gly
19	21	21.9	17	2	S32587	L-ascorbate peroxi
20	21	21.9	17	2	A35550	adrenocortical cel
21	21	21.9	19	2	S59717	hypothetical prote
22	21	21.9	20	2	H49034	nuclear antigen EB
23	20	20.8	15	2	S03955	acidic fibroblast
24	20	20.8	16	2	S01669	rRNA N-glycosidase
25	20	20.8	16	2	S02473	coat protein VP1 -
26	20	20.8	16	2	S70331	endosperm protein,
27	20	20.8	18	2	S45373	translation elonga
28	20	20.8	20	2	S29108	glutathione transf
29	20	20.8	20	2	S63602	glutathione S-tran
30	20	20.8	20	2	PL0192	Ig lambda 2 chain
31	20	20.8	20	2	A44921	hydroxypyruvate re
32	20	20.8	20	2	S46479	retinoid-X-recepto
33	19	19.8	9	3	PC7076	spectrin alpha cha
34	19	19.8	10	2	A13687	caerulein-like pep
35	19	19.8	12	2	S26556	T-cell receptor be
36	19	19.8	13	2	S47365	T-cell antigen rec
37	19	19.8	14	2	S14336	mastoparan B - hor
38	19	19.8	14	2	A47421	leukotriene B-4 12
39	19	19.8	15	2	S36890	ribosomal protein
40	19	19.8	17	2	S61451	hypothetical prote
41	19	19.8	17	2	PH1331	Ig heavy chain DJ
42	19	19.8	18	2	B35910	neurofibromatosis-
43	19	19.8	18	2	S48863	cyclin C - mouse (
44	19	19.8	19	2	S43652	serpin WSZCI - whe
45	19	19.8	19	2	A47689	flagellar sheath p

ALIGNMENTS

RESULT 1
S14560
probable heme-binding protein - garden pea chloroplast (fragment)
C;Species: chloroplast Pisum sativum (garden pea)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jan-2000
C;Accession: S14560
R;Smith, A.G.; Wilson, R.J.; Kaethner, T.M.; Willey, D.L.; Gray, J.C.
submitted to the EMBL Data Library, October 1990
A;Reference number: S14557
A;Accession: S14560
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-18 <SMI>
A;Cross-references: EMBL:X54750; NID:g12194; PID:g12198
C;Genetics:
A;Genome: chloroplast
C;Superfamily: maize chloroplast protein cemA
C;Keywords: chloroplast; heme; transmembrane protein

Query Match 31.2%; Score 30; DB 2; Length 18;
Best Local Similarity 60.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 8 SKKKLFLSLL 17
:||| |: ||
Db 2 AKKKAFIPLL 11

RESULT 2
S13864
methyl coenzyme M reductase (EC 1.8.-.-) II alpha chain - Methanobacterium thermoauto
C;Species: Methanobacterium thermoautotrophicum
A;Variety: strain Marburg
C;Date: 19-Mar-1997 #sequence_revision 13-Sep-1998 #text_change 30-Oct-1998
C;Accession: S13864
R;Rospert, S.; Linder, D.; Ellermann, J.; Thauer, R.K.
Eur. J. Biochem. 194, 871-877, 1990
A;Title: Two genetically distinct methyl-coenzyme M reductases in Methanobacterium th
A;Reference number: S13864; MUID:91099370
A;Accession: S13864
A;Molecule type: protein
A;Residues: 1-14 <ROS>
A;Experimental source: strain Marburg
C;Keywords: methanogenesis; oxidoreductase

Query Match 29.2%; Score 28; DB 2; Length 14;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 9 KKKLFLSLL 17

```
          :|||||  |
Db      3 EKKLFLKAL 11
```

RESULT 3
I52618
hemoglobin beta chain thalassemia mutant Portuguese - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 11-Jul-1996 #text_change 20-Apr-2000
C;Accession: I52618
R;Oner, R.; Oner, C.; Wilson, J.B.; Tamagnini, G.P.; Ribeiro, L.M.; Huisman, T.H.
Br. J. Haematol. 79, 306-310, 1991
A;Title: Dominant beta-thalassaemia trait in a Portuguese family is caused by a deletion
A;Reference number: I52618; MUID:92068764
A;Accession: I52618
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-14 <ONE>
A;Cross-references: GB:S68042; NID:g239717; PIDN:AAB20440.1; PID:g239718
C;Genetics:
A;Gene: GDB:HBB
A;Cross-references: GDB:119297; OMIM:141900
A;Map position: 11p15.4-11p15.4

```
  Query Match             27.1%;  Score 26;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 4.5e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 VQASYSK 9
          |||:| |
Db      3 VQAAYQK 9
```

RESULT 4
A61411
ameletin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 07-May-1999
C;Accession: A61411
R;Burzynski, S.R.
Anal. Biochem. 70, 359-365, 1976
A;Title: Sequential analysis in subnanomolar amounts of peptides. Determination of the s
A;Reference number: A61411; MUID:76182447
A;Accession: A61411
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6 <BUR>
C;Keywords: pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

```
  Query Match             26.0%;  Score 25;  DB 2;  Length 6;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+05;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 QASYSK 9
          || |||
Db      1 QAGYSK 6
```

RESULT 5
I49593
cystic fibrosis transmembrane conductance regulator - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
C;Accession: I49593
R;Denamur, E.; Chehab, F.F.
Hum. Mol. Genet. 3, 1089-1094, 1994
A;Title: Analysis of the mouse and rat CFTR promoter regions.
A;Reference number: I49593; MUID:95072572
A;Accession: I49593
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-17 <RES>
A;Cross-references: GB:L04873; NID:g414726; PIDN:AAA73562.1; PID:g553892
C;Genetics:
A;Gene: CFTR
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding casse

```
  Query Match             25.0%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   41.2%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDF 19
          :| |  :|  |:| | |
Db      1 MQKSPLEKASFISKLFF 17
```

RESULT 6
I84733
gene CFTR protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I84733
R;Denamur, E.; Chehab, F.F.
Hum. Mol. Genet. 3, 1089-1094, 1994
A;Title: Analysis of the mouse and rat CFTR promoter regions.
A;Reference number: I49593; MUID:95072572
A;Accession: I84733
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-17 <RES>
A;Cross-references: GB:L26098; NID:g425185; PIDN:AAA73561.1; PID:g915270
C;Genetics:
A;Gene: CFTR
C;Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding casse

```
  Query Match             25.0%;  Score 24;  DB 2;  Length 17;
  Best Local Similarity   41.2%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDF 19
          :| |  :|  |:| | |
Db      1 MQKSPLEKASFISKLFF 17
```

RESULT 7
A49192
transthyretin - bullfrog (fragment)
N;Alternate names: 3,5,3'-L-triiodothyronine-specific binding protein
C;Species: Rana catesbeiana (bullfrog)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C;Accession: A49192
R;Yamauchi, K.; Kasahara, T.; Hayashi, H.; Horiuchi, R.
Endocrinology 132, 2254-2261, 1993
A;Title: Purification and characterization of a 3,5,3'-L-triiodothyronine-specific bi
A;Reference number: A49192; MUID:93238653
A;Accession: A49192
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-19 <YAM>
A;Experimental source: tadpole plasma
A;Note: sequence extracted from NCBI backbone (NCBIP:130235)

```
  Query Match             24.0%;  Score 23;  DB 2;  Length 19;
  Best Local Similarity   45.5%;  Pred. No. 1.9e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      8 SKKKLFLSLLD 18
          ||  | : :||
Db      7 SKXPLMVKVLD 17
```

```
RESULT   8
S39049
cytotoxin-binding protein - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 18-Feb-1994 #sequence_revision 19-Apr-1996 #text_change 16-Feb-1997
C;Accession: S39049
R;Lutz, F.; Mohr, M.; Grimmig, M.; Leidolf, R.; Linder, D.
Eur. J. Biochem. 217, 1123-1128, 1993
A;Title: Pseudomonas aeruginosa cytotoxin-binding protein in rabbit erythrocyte membrane
A;Reference number: S39049; MUID:94039134
A;Accession: S39049
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <LUT>
C;Superfamily: lens fiber membrane major intrinsic protein

  Query Match             24.0%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 KKKLF 13
          |||:|
Db       5 KKKIF 9


RESULT   9
PH0786
T-cell receptor alpha chain (I7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0786
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0786
A;Molecule type: mRNA
A;Residues: 1-13 <CAS>
A;Cross-references: EMBL:X60889
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

  Query Match             22.9%;  Score 22;  DB 2;  Length 13;
  Best Local Similarity   33.3%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ASYSKKKLF 13
           ::|: | :|
Db       4 SAYANKMIF 12


RESULT  10
S65717
prostaglandin D-synthase - rat (fragment)
N;Alternate names: prostaglandin-H2 D-isomerase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 06-Dec-1996 #sequence_revision 13-Mar-1997 #text_change 31-Oct-1997
C;Accession: S65717
R;Giacomelli, S.; Leone, M.G.; Grima, J.; Silvestrini, B.; Cheng, C.Y.
Biochim. Biophys. Acta 1310, 269-276, 1996
A;Title: Astrocytes synthesize and secrete prostaglandin D synthetase in vitro.
A;Reference number: S65716; MUID:96177373
A;Accession: S65717
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-15 <GIA>
C;Superfamily: lipocalin; lipocalin homology

  Query Match             22.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   40.0%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TVQASYSKKK 11
           ||| :: : |
Db       3 TVQPNFQQDK 12


RESULT  11
PA0046
protein QA100044 - Arabidopsis thaliana (fragment)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 30-Jun-1992 #sequence_revision 06-Jan-1995 #text_change 06-Jun-1997
C;Accession: PA0046; PA0042
R;Kamo, M.; Kawakami, T.; Miyatake, N.; Tsugita, A.
submitted to JIPID, July 1994
A;Description: Separation and characterization of Arabidopsis proteins by two-dimensi
A;Reference number: PA0001
A;Accession: PA0046
A;Molecule type: protein
A;Residues: 1-15 <KAM>
A;Experimental source: stem

  Query Match             22.9%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 2.2e+03;
  Matches    5;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 KKLFLSLLD 18
           || ||:: |
Db       4 KKGFLAVSD 12


RESULT  12
S34444
blaZ protein - Staphylococcus aureus plasmid pI258 (fragment)
C;Species: Staphylococcus aureus
C;Date: 22-Nov-1993 #sequence_revision 09-Mar-1996 #text_change 07-May-1999
C;Accession: S34444
R;Wang, P.Z.; Projan, S.J.; Novick, R.P.
Nucleic Acids Res. 19, 4000, 1991
A;Title: Nucleotide sequence of beta-lactamase regulatory genes from staphylococcal p
A;Reference number: S34444; MUID:91319567
A;Accession: S34444
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-16 <WAN>
A;Cross-references: EMBL:M62650
A;Experimental source: strain RN11
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 199
C;Genetics:
A;Gene: blaZ
A;Genome: plasmid pI258

  Query Match             22.9%;  Score 22;  DB 2;  Length 16;
  Best Local Similarity   50.0%;  Pred. No. 2.4e+03;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      10 KKLFLSLLDF 19
           |:| | :| |
Db       6 KQLQLKILTF 15


RESULT  13
S54272
CTC 75 protein - human
C;Species: Homo sapiens (man)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 17-Mar-1999
C;Accession: S54272
R;Genersch, E.; Eckerskorn, C.; Lottspeich, F.; Herzog, C.; Kuehn, K.; Poeschl, E.
EMBO J. 14, 791-800, 1995
```

A;Title: Purification of the sequence-specific transcription factor CTCBF, involved in t
A;Reference number: S54272; MUID:95188883
A;Accession: S54272
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <GEN>

Query Match 22.9%; Score 22; DB 2; Length 18;
Best Local Similarity 33.3%; Pred. No. 2.6e+03;
Matches 4; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TTVQASYSKKKL 12
           :| :  ||:::|
Db       6 STGKVEYSEEEL 17
```

RESULT 14
G22565
R-phycoerythrin gamma-B chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: G22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: G22565
A;Molecule type: protein
A;Residues: 1-13 <KLO>

Query Match 21.9%; Score 21; DB 2; Length 13;
Best Local Similarity 44.4%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 TTVQASYSK 9
           |  ||:: |
Db       5 TVPQAAFEK 13
```

RESULT 15
S29209
avenin alpha-2 - oat (fragment)
N;Alternate names: CIP-3; coeliac immunoreactive protein 3
C;Species: Avena sativa (oat)
C;Date: 19-Mar-1997 #sequence_revision 24-Jul-1998 #text_change 24-Jul-1998
C;Accession: S29209
R;Rocher, A.; Colilla, F.; Ortiz, M.L.; Mendez, E.
FEBS Lett. 310, 37-40, 1992
A;Title: Identification of the three major coeliac immunoreactive proteins and one alpha
A;Reference number: S29207; MUID:92405739
A;Accession: S29209
A;Molecule type: protein
A;Residues: 1-14 <ROC>
A;Experimental source: endosperm
C;Superfamily: gliadin
C;Keywords: prolamin; seed

Query Match 21.9%; Score 21; DB 2; Length 14;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 TTVQASYSKK 10
           |||| : |::
Db       2 TTVQYNPSEQ 11
```

Search completed: December 21, 2000, 08:30:08
Job time: 272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:33:00 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-35
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	24	25.0	18	1	UC21_MAIZE	P80627	zea mays (m
2	23	24.0	12	1	PA21_MICFM	P25072	micrurus fu
3	23	24.0	20	1	CPBX_CAVPO	P34033	cavia porce
4	22	22.9	20	1	VR90_BORPE	P81549	bordetella
5	21	21.9	10	1	XYNB_DICB4	P80717	dictyoglomu
6	21	21.9	20	1	MI17_BOVIN	P35451	bos taurus
7	21	21.9	20	1	TL18_SPIOL	P82536	spinacia ol
8	20	20.8	12	1	PA2B_VIPBO	P31859	vipera beru
9	20	20.8	15	1	FGF1_CANFA	P18651	canis famil
10	20	20.8	16	1	RIPK_TRIKI	P16093	trichosanth
11	19	19.8	8	1	B44K_PORGI	P81886	porphyromon
12	19	19.8	14	1	MAST_VESBA	P21654	vespa basal
13	19	19.8	15	1	EF1A_MICCR	P81266	microplitis
14	19	19.8	16	1	RL6_VIBPR	Q56715	vibrio prot
15	19	19.8	19	1	PHSL_DESBN	P13066	desulfovibr
16	19	19.8	20	1	STYA_STYCL	P81469	styela clav
17	18	18.8	9	1	ULAE_HUMAN	P31931	homo sapien
18	18	18.8	13	1	UVRD_SALTY	Q05311	salmonella
19	18	18.8	13	1	YPE2_LACLC	P42021	lactococcus
20	18	18.8	16	1	ARCD_PSEPU	P41147	pseudomonas
21	18	18.8	17	1	ATPI_PAVLU	P28529	pavlova lut
22	18	18.8	19	1	UP24_UPEIN	P82030	uperoleia i
23	18	18.8	20	1	AMP_FUSNU	P81207	fusobacteri
24	18	18.8	20	1	CRP_MUSCA	P19094	mustelus ca
25	17	17.7	8	1	CCKN_MACEU	P30369	macropus eu
26	17	17.7	11	1	TKN_ELEMO	P01293	eledone mos
27	17	17.7	12	1	NO40_SOYBN	P55960	glycine max
28	17	17.7	14	1	MAST_VESMA	P04205	vespa manda
29	17	17.7	14	1	MAST_VESXA	P01515	vespa xanth
30	17	17.7	14	1	UC34_MAIZE	P80640	zea mays (m
31	17	17.7	15	1	CBPB_PROAT	P19628	protopterus
32	17	17.7	17	1	BOL1_MEGPE	P10521	megabombus
33	17	17.7	18	1	NODD_RHILO	Q52838	rhizobium l
34	17	17.7	18	1	RL23_HALCU	P05975	halobacteri
35	17	17.7	20	1	CATA_ACIRA	P81422	acinetobact
36	17	17.7	20	1	RIPX_CUCPE	P80750	cucurbita p
37	16	16.7	8	1	AKH_MELML	P25423	melolontha
38	16	16.7	8	1	COXG_RAT	P80430	rattus norv
39	16	16.7	8	1	CPD1_ENTFA	P13269	enterococcu
40	16	16.7	10	1	COXM_RAT	P80431	rattus norv
41	16	16.7	10	1	PNEU_RAT	P21996	rattus norv
42	16	16.7	11	1	CSI5_BACSU	P81095	bacillus su
43	16	16.7	11	1	UXB2_YEAST	P99013	saccharomyc
44	16	16.7	14	1	MAST_VESLE	P01514	vespula lew
45	16	16.7	14	1	RS19_PPWBP	Q52093	pigeon pea

ALIGNMENTS

RESULT 1
UC21_MAIZE
ID UC21_MAIZE STANDARD; PRT; 18 AA.
AC P80627;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE UNKNOWN PROTEIN FROM 2D-PAGE OF ETIOLATED COLEOPTILE (SPOT 443)
DE (FRAGMENT).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN [1]
RP SEQUENCE.
RC TISSUE=COLEOPTILE;
RA Touzet P., Riccardi F., Morin C., Damerval C., Huet J.-C.,
RA Pernollet J.-C., Zivy M., de Vienne D.;
RT "The maize two dimensional gel protein database: towards an integrated
RT genome analysis program.";
RL Theor. Appl. Genet. 93:997-1005(1996).
CC -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC PROTEIN IS: 5.9, ITS MW IS: 26.5 KDA.
CC -!- SIMILARITY: TO L-ASCORBATE PEROXIDASES.
DR HSSP; P48534; 1APX.
DR MAIZE-2DPAGE; P80627; COLEOPTILE.
DR MAIZEDB; 123953; -.
FT NON_TER 1 1
FT NON_TER 18 18
SQ SEQUENCE 18 AA; 1938 MW; F32F6FEF038BAB8A CRC64;

Query Match 25.0%; Score 24; DB 1; Length 18;
Best Local Similarity 71.4%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 TVQASYS 8
          || | ||
Db      6 TVSAEYS 12
```

RESULT 2
PA21_MICFM
ID PA21_MICFM STANDARD; PRT; 12 AA.
AC P25072;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE PHOSPHOLIPASE A2 ISOZYME 1 (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE 2-ACYLHYDROLASE) (FRAGMENT).
OS Micrurus fulvius microgalbineus (Mexican coral snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Elapinae; Micrurus.
RN [1]
RP SEQUENCE.
RC TISSUE=VENOM;

```
RX   MEDLINE; 79255521.
RA   Possani L.D., Alagon A.C., Fletcher P.L. Jr., Varela M.J., Julia J.Z.;
RT   "Purification and characterization of a phospholipase A2 from the
RT   venom of the coral snake, Micrurus fulvius microgalbineus (Brown and
RT   Smith).";
RL   Biochem. J. 179:603-606(1979).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   INTERPRO; IPR001211; -.
DR   PROSITE; PS00118; PA2_HIS; PARTIAL.
DR   PROSITE; PS00119; PA2_ASP; PARTIAL.
KW   Hydrolase; Lipid degradation; Calcium; Venom.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1398 MW;  CC21992A899F0339 CRC64;


  Query Match             24.0%;  Score 23;  DB 1;  Length 12;
  Best Local Similarity   66.7%;  Pred. No. 3.6e+02;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     15 SLLDFQ 20
          |||:|:
Db      1 SLLBFK 6


RESULT   3
CPBX_CAVPO
ID   CPBX_CAVPO     STANDARD;      PRT;    20 AA.
AC   P34033;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CYTOCHROME P450IIB (EC 1.14.14.1) (FRAGMENT).
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
RN   [1]
RP   SEQUENCE, AND CHARACTERIZATION.
RC   STRAIN=HARTLEY; TISSUE=LIVER;
RX   MEDLINE; 91054472.
RA   Narimatsu S., Akutsu Y., Matsunaga T., Watanabe K., Yamamoto I.,
RA   Yoshimura H.;
RT   "Purification of a cytochrome P450 isozyme belonging to a subfamily
RT   of P450 IIB from liver microsomes of guinea pigs.";
RL   Biochem. Biophys. Res. Commun. 172:607-613(1990).
CC   -!- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE
CC       MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN
CC       NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. THIS ISOZYME IS ACTIVE
CC       UPON P.NITROANISOLE, ANILINE, D-BENZPHETAMINE, DELTA(9)-
CC       TETRAHYDROCANNABINOL (THC) AND STRYCHNINE.
CC   -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +
CC       OXIDIZED FLAVOPROTEIN + H(2)O.
CC   -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR   PIR; A36154; A36154.
DR   INTERPRO; IPR001128; -.
DR   PROSITE; PS00086; CYTOCHROME_P450; PARTIAL.
KW   Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW   Microsome; Endoplasmic reticulum.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2259 MW;  78DC81280C970A55 CRC64;


  Query Match             24.0%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.3e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     12 LFLSLL 17
          |||:||
Db      7 LFLALL 12


RESULT   4
VR90_BORPE
ID   VR90_BORPE     STANDARD;      PRT;    20 AA.
AC   P81549;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   VIRULENCE-ASSOCIATED OUTER MEMBRANE PROTEIN VIR90 (FRAGMENT).
GN   VIR90.
OS   Bordetella pertussis.
OC   Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
OC   Bordetella.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=TOHAMA I;
RX   MEDLINE; 99179239.
RA   Passerini de Rossi B.N., Friedman L.E., Gonzalez Flecha F.L.,
RA   Castello P.R., Franco M.A., Rossi J.P.F.C.;
RT   "Identification of Bordetella pertussis virulence-associated outer
RT   membrane proteins.";
RL   FEMS Microbiol. Lett. 172:9-13(1999).
KW   Outer membrane; Virulence.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2116 MW;  D24E1CDCA665206C CRC64;


  Query Match             22.9%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 9.4e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 TVQASYS 8
          ||:  ||
Db     11 TVEGEYS 17


RESULT   5
XYNB_DICB4
ID   XYNB_DICB4     STANDARD;      PRT;    10 AA.
AC   P80717;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   ENDO-1,4-BETA-XYLANASE B (EC 3.2.1.8) (XYLANASE B)
DE   (1,4-BETA-D-XYLAN XYLANOHYDROLASE B) (FRAGMENT).
OS   Dictyoglomus sp. (strain B4A).
OC   Bacteria; Dictyoglomus group; Dictyoglomus.
RN   [1]
RP   SEQUENCE.
RA   Adamsen A.K., Jacobsen S., Ahring B.K.;
RL   Submitted (OCT-1996) to the SWISS-PROT data bank.
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-BETA-D-XYLOSIDIC
CC       LINKAGES IN XYLANS.
CC   -!- PATHWAY: XYLAN DEGRADATION.
CC   -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC       HYDROLASES).
DR   INTERPRO; IPR001000; -.
DR   PROSITE; PS00591; GLYCOSYL_HYDROL_F10; PARTIAL.
KW   Xylan degradation; Hydrolase; Glycosidase.
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1144 MW;  4554322AA72041A3 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   55.6%;  Pred. No. 6.8e+02;
  Matches    5;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy     10 KKLFLSLLD 18
          ||  | | |
Db      2 KKTILDLKD 10
```

```
RESULT   6
MI17_BOVIN
ID   MI17_BOVIN     STANDARD;      PRT;    20 AA.
AC   P35451;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   17 KDA MILK GLYCOPROTEIN (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=MILK;
RX   MEDLINE; 93308294.
RA   Soerensen E.S., Petersen T.E.;
RT   "Purification and characterization of three proteins isolated from
RT   the proteose peptone fraction of bovine milk.";
RL   J. Dairy Res. 60:189-197(1993).
CC   -!- PTM: N-GLYCOSYLATED.
CC   -!- SIMILARITY: TO CAMEL WHEY PROTEIN.
KW   Glycoprotein; Milk.
FT   NON_TER       1      1
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2233 MW;  4CCA589404C62C27 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   35.3%;  Pred. No. 1.4e+03;
  Matches    6;  Conservative    3;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSKKKLFLSLL 17
          :: |      || ||:|
Db      1 SSXQPQSQNPKLPLSIL 17


RESULT   7
TL18_SPIOL
ID   TL18_SPIOL     STANDARD;      PRT;    20 AA.
AC   P82536;
DT   01-OCT-2000 (Rel. 40, Created)
DT   01-OCT-2000 (Rel. 40, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   THYLAKOID LUMENAL 18 KDA PROTEIN (P18) (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RA   Kieselbach T., Bystedt M., Schroeder W.P.;
RL   Submitted (MAY-2000) to the SWISS-PROT data bank.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE LUMEN.
KW   Chloroplast; Thylakoid membrane.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2192 MW;  752C21963F49FA64 CRC64;


  Query Match             21.9%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   30.8%;  Pred. No. 1.4e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSKKKLF 13
          | :|:  : | :|
Db      5 TPLQSKVTNKVVF 17


RESULT   8
PA2B_VIPBO
ID   PA2B_VIPBO     STANDARD;      PRT;    12 AA.
AC   P31859;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PHOSPHOLIPASE A2, BASIC (EC 3.1.1.4) (PHOSPHATIDYLCHOLINE
DE   2-ACYLHYDROLASE) (FRAGMENT).
OS   Vipera berus orientalis (Viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Vipera.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 74128698.
RA   Delori P.J.;
RT   "Purification and physiochemical, chemical and biological properties
RT   of a toxic A2 phospholipase isolated from the venom of viperidae
RT   snakes: Vipera berus.";
RL   Biochimie 55:1031-1045(1973).
CC   -!- FUNCTION: PA2 CATALYZES THE CALCIUM-DEPENDENT HYDROLYSIS OF THE
CC       2-ACYL GROUPS IN 3-SN-PHOSPHOGLYCERIDES.
CC   -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + H(2)O = 1-ACYLGLYCERYL-
CC       PHOSPHOCHOLINE + A FATTY ACID ANION.
CC   -!- SIMILARITY: BELONGS TO THE PHOSPHOLIPASE A2 FAMILY.
DR   INTERPRO; IPR001211; -.
DR   PROSITE; PS00118; PA2_HIS; PARTIAL.
DR   PROSITE; PS00119; PA2_ASP; PARTIAL.
KW   Hydrolase; Lipid degradation; Calcium; Venom.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1412 MW;  D054351A89969879 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 12;
  Best Local Similarity   80.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     15 SLLDF 19
            |||:|
Db      1 SLLEF 5


RESULT   9
FGF1_CANFA
ID   FGF1_CANFA     STANDARD;      PRT;    15 AA.
AC   P18651;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   HEPARIN-BINDING GROWTH FACTOR 1 (HBGF-1) (ACIDIC FIBROBLAST GROWTH
DE   FACTOR) (AFGF) (ALPHA-ENDOTHELIAL CELL GROWTH FACTOR) (FRAGMENT).
GN   FGF1 OR FGF-1.
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89231704.
RA   Quinkler W., Maasberg M., Bernotat-Danielowski S., Luethe N.,
RA   Sharma H.S., Schaper W.;
RT   "Isolation of heparin-binding growth factors from bovine, porcine and
RT   canine hearts.";
RL   Eur. J. Biochem. 181:67-73(1989).
CC   -!- FUNCTION: THE HEPARIN-BINDING GROWTH FACTORS ARE ANGIOGENIC AGENTS
CC       IN VIVO AND ARE POTENT MITOGENS FOR A VARIETY OF CELL TYPES IN
CC       VITRO. THERE ARE DIFFERENCES IN THE TISSUE DISTRIBUTION AND
CC       CONCENTRATION OF THESE 2 GROWTH FACTORS.
CC   -!- SUBUNIT: MONOMER.
CC   -!- MISCELLANEOUS: THIS PROTEIN BINDS HEPARIN, ALTHOUGH LESS STRONGLY
CC       THAN DOES BFGF.
CC   -!- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY.
DR   PIR; S03955; S03955.
DR   HSSP; P05230; 2AXM.
DR   INTERPRO; IPR002209; -.
DR   PROSITE; PS00247; HBGF_FGF; PARTIAL.
KW   Growth factor; Mitogen; Vascularization; Heparin-binding.
```

```
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1732 MW;  53CC9A3CADDDDAA1 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   57.1%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 SYSKKKL 12
          :| | ||
Db      1 NYMKPKL 7


RESULT  10
RIPK_TRIKI
ID   RIPK_TRIKI     STANDARD;      PRT;    16 AA.
AC   P16093;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   RIBOSOME-INACTIVATING PROTEIN TRICHOKIRIN (RRNA N-GLYCOSIDASE)
DE   (EC 3.2.2.22) (FRAGMENT).
OS   Trichosanthes kirilowii (Mongolian snake-gourd).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Cucurbitales; Cucurbitaceae; Trichosanthes.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SEED;
RX   MEDLINE; 89005108.
RA   Casellas P., Dussossoy D., Falasca A.I., Barbieri L.,
RA   Guillemot J.C., Ferrara P., Bolognesi A., Cenini P., Stirpe F.;
RT   "Trichokirin, a ribosome-inactivating protein from the seeds of
RT   Trichosanthes kirilowii Maximowicz. Purification, partial
RT   characterization and use for preparation of immunotoxins.";
RL   Eur. J. Biochem. 176:581-588(1988).
CC   -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC       SPECIFIC ADENOSINE ON THE 28S RRNA.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING
CC       PROTEINS. BELONGS TO TYPE 1 RIP.
DR   PIR; S01669; S01669.
DR   INTERPRO; IPR001574; -.
DR   PROSITE; PS00275; SHIGA_RICIN; PARTIAL.
KW   Protein synthesis inhibitor; Hydrolase; Toxin; Glycoprotein.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1605 MW;  5E268A7F345935A2 CRC64;


  Query Match             20.8%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 ASYSK 9
          ||| |
Db     12 ASYEK 16


RESULT  11
B44K_PORGI
ID   B44K_PORGI     STANDARD;      PRT;     8 AA.
AC   P81886;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   44 KDA IMMUNOGENIC PROTEIN (FRAGMENT).
OS   Porphyromonas gingivalis (Bacteroides gingivalis).
OC   Bacteria; CFB group; Bacteroidaceae; Porphyromonas.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=VPB 3492;
RX   MEDLINE; 20198497.
RA   Norris J.M., Love D.N.;
RT   "Serum antibody responses of cats to soluble whole cell antigens of
RT   feline porphyromonas gingivalis.";
RL   Vet. Microbiol. 73:37-49(2000).
CC   -!- SIMILARITY: TO P.GINGIVALIS HEMAGGLUTININ A.
KW   Antigen.
FT   NON_TER      8      8
SQ   SEQUENCE   8 AA;  989 MW;  9554540326CB476D CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      5 ASYSKKKL 12
            | | |: :
Db      1 APYQKRNI 8


RESULT  12
MAST_VESBA
ID   MAST_VESBA     STANDARD;      PRT;    14 AA.
AC   P21654;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-MAY-1991 (Rel. 18, Last sequence update)
DT   01-AUG-1992 (Rel. 23, Last annotation update)
DE   MASTOPARAN B.
OS   Vespa basalis (Hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RX   MEDLINE; 91174755.
RA   Lo C.-L., Hwang L.-L.;
RT   "Structure and biological activities of a new mastoparan isolated
RT   from the venom of the hornet Vespa basalis.";
RL   Biochem. J. 274:453-456(1991).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. ACTIVATES G PROTEINS
CC       THAT COUPLE TO PHOSPHOLIPASE C.
DR   PIR; S14336; S14336.
KW   Mast cell degranulation; Venom; Amidation.
FT   MOD_RES      14     14       AMIDATION.
SQ   SEQUENCE   14 AA;  1613 MW;  D35944CA193A19A2 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 14;
  Best Local Similarity   57.1%;  Pred. No. 2.2e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 SYSKKKL 12
          |::|| |
Db      8 SWAKKVL 14


RESULT  13
EF1A_MICCR
ID   EF1A_MICCR     STANDARD;      PRT;    15 AA.
AC   P81266;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   ELONGATION FACTOR 1-ALPHA (EF-1-ALPHA) (FRAGMENT).
OS   Microplitis croceipes.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC   Ichneumonoidea; Braconidae; Microgastrinae; Microplitis.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 99033459.
RA   Stuart M.K.;
RT   "An antibody diagnostic for hymenopteran parasitism is specific for a
RT   homologue of elongation factor-1 alpha.";
```

```
RL   Arch. Insect Biochem. Physiol. 39:1-8(1998).
CC   -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT BINDING OF
CC       AMINOACYL-TRNA TO THE A-SITE OF RIBOSOMES DURING PROTEIN
CC       BIOSYNTHESIS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC       EF-TU/EF-1A SUBFAMILY.
DR   INTERPRO; IPR000795; -.
DR   PROSITE; PS00301; EFACTOR_GTP; PARTIAL.
KW   Elongation factor; Protein biosynthesis; GTP-binding.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1670 MW;  1EB7DA05B09B3751 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 15;
  Best Local Similarity   20.0%;  Pred. No. 2.3e+03;
  Matches    2;  Conservative    7;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 SKKKLFLSLL 17
           :|:|: ::::
Db       1 AKEKIHINIV 10


RESULT  14
RL6_VIBPR
ID   RL6_VIBPR      STANDARD;      PRT;    16 AA.
AC   Q56715;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   50S RIBOSOMAL PROTEIN L6 (FRAGMENT).
GN   RPLF.
OS   Vibrio proteolyticus (Aeromonas proteolytica).
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15338;
RX   MEDLINE; 97149305.
RA   Setterquist R.A., Smith G.K., Oakley T.H., Lee Y.H., Fox G.E.;
RT   "Sequence, overproduction and purification of Vibrio proteolyticus
RT   ribosomal protein L18 for in vitro and in vivo studies.";
RL   Gene 183:237-242(1996).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS
CC       LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE
CC       PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U38943; AAB41328.1; -.
DR   INTERPRO; IPR002358; -.
DR   PROSITE; PS00525; RIBOSOMAL_L6_1; PARTIAL.
KW   Ribosomal protein; rRNA-binding.
FT   NON_TER      1      1
SQ   SEQUENCE   16 AA;  1935 MW;  ABC19078DF581B6C CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 16;
  Best Local Similarity   44.4%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKK 11
           |:   :|||
Db       8 VRTKEAKKK 16


RESULT  15
PHSL_DESBN
ID   PHSL_DESBN     STANDARD;      PRT;    19 AA.
AC   P13066;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   PERIPLASMIC [NIFESE] HYDROGENASE LARGE SUBUNIT (EC 1.18.99.1) (NIFESE
DE   HYDROGENLYASE LARGE CHAIN) (FRAGMENT).
OS   Desulfovibrio baculatus (strain Norway 4).
OC   Bacteria; Proteobacteria; delta subdivision; Desulfomicrobium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88106446.
RA   Prickril B.C., He S.H., Li C., Menon N., Choi E.S., Przybyla A.E.,
RA   Dervartanian D.V., Peck H.D. Jr., Fauque G., le Gall J., Teixeira M.,
RA   Moura I., Moura J.J.G., Patil D., Huynh B.H.;
RT   "Identification of three classes of hydrogenase in the genus,
RT   Desulfovibrio.";
RL   Biochem. Biophys. Res. Commun. 149:369-377(1987).
CC   -!- CATALYTIC ACTIVITY: 2 REDUCED FERREDOXIN + 2 H(+) = 2 OXIDIZED
CC       FERREDOXIN + H(2).
CC   -!- COFACTOR: ONE NICKEL ION, TWO 4FE-4S CLUSTERS AND ONE
CC       SELENOCYSTEINE.
CC   -!- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
CC   -!- SUBCELLULAR LOCATION: PERIPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE [NIFE]/[NIFESE] HYDROGENASE LARGE
CC       SUBUNIT FAMILY.
DR   PIR; H27480; H27480.
DR   INTERPRO; IPR001501; -.
DR   PROSITE; PS00507; NI_HGENASE_L_1; PARTIAL.
DR   PROSITE; PS00508; NI_HGENASE_L_2; PARTIAL.
KW   Oxidoreductase; Periplasmic; Nickel; Selenium; Selenocysteine.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  1942 MW;  2BFCD2D360F00367 CRC64;


  Query Match             19.8%;  Score 19;  DB 1;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 3e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 ASYSKKKL 12
           |:  |||:
Db       7 AADGKKKI 14


Search completed: December 21, 2000, 08:33:00
Job time: 443 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:40 ; Search time 57.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-35
Perfect score: 96
Sequence: 1 TTVQASYSKKKLFLSLLDFQ 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	27.1	18	3	Q02414	Q02414 aspergillus
2	26	27.1	19	2	Q9RLU4	Q9rlu4 lactococcus
3	25	26.0	18	11	P97522	P97522 rattus norv
4	24	25.0	15	3	Q9URE0	Q9ure0 saccharomyc
5	24	25.0	15	4	Q9UCH4	Q9uch4 homo sapien
6	24	25.0	18	2	O06514	O06514 escherichia
7	24	25.0	20	2	Q9R4E9	Q9r4e9 rhodopseudo
8	23	24.0	17	2	P78199	P78199 escherichia
9	23	24.0	18	2	Q56610	Q56610 vibrio chol
10	23	24.0	18	11	Q9QVB0	Q9qvb0 rattus sp.
11	23	24.0	20	6	Q9TS18	Q9ts18 oryctolagus
12	23	24.0	20	8	P82138	P82138 spinacia ol
13	23	24.0	20	11	O08806	O08806 mus musculu
14	22	22.9	15	4	P78533	P78533 homo sapien
15	22	22.9	16	4	Q9UD41	Q9ud41 homo sapien
16	22	22.9	16	6	Q9TRA2	Q9tra2 sus scrofa
17	22	22.9	17	11	Q9Z170	Q9z170 mus musculu
18	22	22.9	19	4	Q9UMM9	Q9umm9 homo sapien
19	22	22.9	19	6	Q9TRF3	Q9trf3 macropus gi
20	21	21.9	9	2	Q45852	Q45852 clostridium
21	21	21.9	12	11	O54970	O54970 mus musculu
22	21	21.9	14	2	Q9R5Q6	Q9r5q6 burkholderi
23	21	21.9	18	11	Q9QUZ8	Q9quz8 rattus sp.
24	21	21.9	19	3	Q92321	Q92321 saccharomyc
25	21	21.9	20	2	Q49448	Q49448 mycoplasma
26	21	21.9	20	9	Q37868	Q37868 bacteriopha
27	21	21.9	20	12	Q86935	Q86935 human herpe
28	21	21.9	20	12	Q86936	Q86936 human herpe
29	21	21.9	20	12	Q86937	Q86937 human herpe
30	21	21.9	20	12	Q86938	Q86938 human herpe
31	21	21.9	20	12	Q86939	Q86939 human herpe
32	21	21.9	20	12	Q86940	Q86940 human herpe
33	21	21.9	20	12	Q86941	Q86941 human herpe
34	21	21.9	20	12	Q86942	Q86942 human herpe
35	20.5	21.4	19	13	Q9PRP1	Q9prp1 gallus gall
36	20	20.8	13	11	Q62354	Q62354 mus musculu
37	20	20.8	13	11	Q62355	Q62355 mus musculu
38	20	20.8	15	2	O68425	O68425 buchnera ap
39	20	20.8	15	2	Q9R4Y0	Q9r4y0 pseudomonas
40	20	20.8	16	2	Q47605	Q47605 escherichia
41	20	20.8	16	2	Q9R4L0	Q9r4l0 spiroplasma
42	20	20.8	16	10	Q9S8A4	Q9s8a4 secale cere
43	20	20.8	16	12	Q79458	Q79458 human immun
44	20	20.8	17	8	Q9ZY77	Q9zy77 trissolcus
45	20	20.8	18	2	Q46390	Q46390 clostridium

ALIGNMENTS

RESULT 1
Q02414
ID Q02414 PRELIMINARY; PRT; 18 AA.
AC Q02414;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE OROTIDINE-5'-MONOPHOSPHATE DECARBOXYLASE (FRAGMENT).
GN PYRG.
OS Aspergillus awamori.
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae;
OC anamorphic Trichocomaceae; Aspergillus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96031709.
RA Gouka R.J., Hessing J.G., Stam H., Musters W., Hondel C.A.;
RT "A novel strategy for the isolation of defined pyrG mutants and the
RT development of a site-specific integration system for Aspergillus
RT awamori.";
RL Curr. Genet. 27:536-540(1995).
DR EMBL; S79674; AAB35350.1; -.
FT NON_TER 1 1
SQ SEQUENCE 18 AA; 1891 MW; E1A0E34C7D23688E CRC64;

Query Match 27.1%; Score 26; DB 3; Length 18;
Best Local Similarity 41.2%; Pred. No. 8.1e+02;
Matches 7; Conservative 1; Mismatches 9; Indels 0; Gaps 0;

```
Qy       2 TVQASYSKKKLFLSLLD 18
           || |  :  |  | | |
Db       2 TVSADVTTTKELLDLAD 18
```

RESULT 2
Q9RLU4
ID Q9RLU4 PRELIMINARY; PRT; 19 AA.
AC Q9RLU4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE NISG PROTEIN (FRAGMENT).

```
GN   NISG.
OS   Lactococcus lactis.
OG   Plasmid pLEB513.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Wahlstroem G., Takala T., Saris P.E.J.;
RT   "Evidence for a mosaic structure of the Tn5481 in Lactococcus lactis
RT   N8.";
RL   DNA Seq. 9:245-261(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N8; TRANSPOSON=TN5481;
RA   Immonen T., Saris P.E.J.;
RT   "Characterization of the nisFEG operon of the nisin Z producing
RT   Lactococcus lactis subsp. lactis N8 strain.";
RL   DNA Seq. 9:263-274(1998).
DR   EMBL; AJ000993; CAA04441.1; -.
KW   Plasmid.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2139 MW;  6A429DC80E673613 CRC64;


  Query Match             27.1%;  Score 26;  DB 2;  Length 19;
  Best Local Similarity   41.7%;  Pred. No. 8.6e+02;
  Matches    5;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TTVQASYSKKKL 12
          | :   : ||||:
Db      7 TIILVALSKKKI 18


RESULT   3
P97522
ID   P97522      PRELIMINARY;      PRT;    18 AA.
AC   P97522;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   CFTR GENE.
GN   CFTR.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Vuillaumier S., Denamur E.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X95927; CAA65168.1; -.
SQ   SEQUENCE   18 AA;  2158 MW;  5C5855056C1CE6DE CRC64;


  Query Match             26.0%;  Score 25;  DB 11;  Length 18;
  Best Local Similarity   38.9%;  Pred. No. 1.2e+03;
  Matches    7;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFLSLLDFQ 20
          :| |  :|  |:| | |:
Db      1 MQKSPLEKASFISKLFFR 18


RESULT   4
Q9URE0
ID   Q9URE0      PRELIMINARY;      PRT;    15 AA.
AC   Q9URE0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   TYPE II TOPOISOMERASE, TOPOISOMERASE II.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Elsea S.H., Hsiung Y., Nitiss J.L., Osheroff N.;
RL   J. Biol. Chem. 270:1913-1920(1995).
SQ   SEQUENCE   15 AA;  2094 MW;  0A6A37F6E81E85F6 CRC64;


  Query Match             25.0%;  Score 24;  DB 3;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 1.5e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      7 YSKKKLFLS 15
          | |:| ::|
Db      2 YQKRKDYMS 10


RESULT   5
Q9UCH4
ID   Q9UCH4      PRELIMINARY;      PRT;    15 AA.
AC   Q9UCH4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NATURAL KILLER ENHANCING FACTOR, NKEF.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93215001.
RA   Shau H., Gupta R.K., Golub S.H.;
RT   "Identification of a natural killer enhancing factor (NKEF) from human
RT   erythroid cells.";
RL   Cell. Immunol. 147:1-11(1993).
SQ   SEQUENCE   15 AA;  1738 MW;  4909D4A793D382BF CRC64;


  Query Match             25.0%;  Score 24;  DB 4;  Length 15;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+03;
  Matches    6;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      3 VQASYSKKKLFL 14
          |||   |  :||
Db      2 VQAFQGKVNVFL 13


RESULT   6
O06514
ID   O06514      PRELIMINARY;      PRT;    18 AA.
AC   O06514;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUL-1997 (TrEMBLrel. 04, Last annotation update)
DE   DNA, MOSAIC MERCURY RESISTANCE TRANSPOSABLE ELEMENT
DE   (MER-OPERON)1040 BP (FRAGMENT).
GN   MERP.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CH210; TRANSPOSON=TN5059;
RX   MEDLINE; 97303088.
RA   Nikiforov V., Kholodii G., Minakhin L., Gorlenko Z., Kalyaeva E.,
RA   Mindlin S., Nikiforov V.;
RT   "Intercontinental spread of promiscuous mercury-resistance transposons
RT   in environmental bacteria.";
RL   Mol. Microbiol. 24:321-329(1997).
DR   EMBL; Y09026; CAA70241.1; -.
FT   NON_TER       18     18
SQ   SEQUENCE   18 AA;  1899 MW;  484A2D7197814DF9 CRC64;
```

```
  Query Match             25.0%;  Score 24;  DB 2;  Length 18;
  Best Local Similarity   71.4%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 KKLFLSL 16
           |||| :|
Db       2 KKLFAAL 8


RESULT   7
Q9R4E9
ID   Q9R4E9      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4E9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   COA THIOESTERASE (FRAGMENT).
OS   Rhodopseudomonas palustris.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Rhodopseudomonas.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96171228.
RA   Kuver J., Xu Y., Gibson J.;
RT   "Metabolism of cyclohexane carboxylic acid by the photosynthetic
RT   bacterium Rhodopseudomonas palustris.";
RL   Arch. Microbiol. 164:337-345(1995).
SQ   SEQUENCE   20 AA;  2259 MW;  870D483384D44315 CRC64;


  Query Match             25.0%;  Score 24;  DB 2;  Length 20;
  Best Local Similarity   46.2%;  Pred. No. 2e+03;
  Matches    6;  Conservative    3;  Mismatches    2;  Indels    2;  Gaps    1;

Qy      10 KKLF--LSLLDFQ 20
           | ||  :|:|| :
Db       2 KSLFDLISILDLE 14


RESULT   8
P78199
ID   P78199      PRELIMINARY;      PRT;    17 AA.
AC   P78199;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   PROBABLE N-ACETYLMURAMOYL-L-ALANINE AMIDASE (EC 3.5.1.28) (FRAGMENT).
GN   AMIA.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE; 97349980.
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).
DR   EMBL; D90874; BAA16324.1; -.
FT   NON_TER      1      1
SQ   SEQUENCE   17 AA;  2192 MW;  02EDC1B821D3431D CRC64;


  Query Match             24.0%;  Score 23;  DB 2;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.5e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 QASYSKKK 11
            | ::|||:
Db      10 QKAHSKKR 17


RESULT   9
Q56610
ID   Q56610      PRELIMINARY;      PRT;    18 AA.
AC   Q56610;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ACCA (FRAGMENT).
GN   ACCA.
OS   Vibrio cholerae.
OC   Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C6706;
RX   MEDLINE; 97074686.
RA   Franco A., Peir-En Y., Johnson J., Barry E.M., Guerra H., Maurer R.,
RA   Morris J.G.;
RT   "Cloning and characterization of dnaE, encoding the catalytic subunit
RT   of replicative DNA polymerase III, from Vibrio cholerae strain
RT   C6706.";
RL   Gene 175:281-283(1996).
DR   EMBL; U30472; AAC44579.1; -.
FT   NON_TER      18     18
SQ   SEQUENCE   18 AA;  2153 MW;  18EBCDAD212842EF CRC64;


  Query Match             24.0%;  Score 23;  DB 2;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 LSLLDFQ 20
           |: |||:
Db       3 LNFLDFE 9


RESULT  10
Q9QVB0
ID   Q9QVB0      PRELIMINARY;      PRT;    18 AA.
AC   Q9QVB0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   (CA2+-MG2+)-ATPASE (FRAGMENT).
OS   Rattus sp.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93054417.
RA   Ezaki J., Himeno M., Kato K.;
RT   "Purification and characterization of (Ca2+-Mg2+)-ATPase in rat liver
RT   lysosomal membranes.";
RL   J. Biochem. 112:33-39(1992).
SQ   SEQUENCE   18 AA;  2008 MW;  8314BA373ED6AA47 CRC64;


  Query Match             24.0%;  Score 23;  DB 11;  Length 18;
  Best Local Similarity   44.4%;  Pred. No. 2.6e+03;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 KKKLFLSLL 17
            |.:::|: ||
Db       6 KERVFMVLL 14


RESULT  11
```

```
Q9TS18
ID   Q9TS18      PRELIMINARY;      PRT;    20 AA.
AC   Q9TS18;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOTOXIN-BINDING PROTEIN (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 94039134.
RA   Lutz F., Mohr M., Grimmig M., Leidolf R., Linder D.;
RT   "Pseudomonas aeruginosa cytotoxin-binding protein in rabbit
RT   erythrocyte membranes. An oligomer of 28 kDa with similarity to
RT   transmembrane channel proteins.";
RL   Eur. J. Biochem. 217:1123-1128(1993).
SQ   SEQUENCE   20 AA;  2414 MW;  40B82D7DB5283D2D CRC64;

  Query Match             24.0%;  Score 23;  DB 6;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 KKKLF 13
           |||:|
Db       5 KKKIF 9


RESULT  12
P82138
ID   P82138      PRELIMINARY;      PRT;    20 AA.
AC   P82138;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CHLOROPLAST 30S RIBOSOMAL PROTEIN S15 (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Caryophyllidae; Caryophyllales;
OC   Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. ALWARO; TISSUE=LEAF;
RA   Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT   "Identification of all the proteins in the small subunit of an
RT   organelle (chloroplast) ribosome.";
RL   J. Biol. Chem. 0:0-0(2000).
CC   -!- FUNCTION: THIS PROTEINS BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST AND OTHER PLASTIDS.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- SIMILARITY: BELONGS TO THE S15P FAMILY OF RIBOSOMAL PROTEINS.
DR   INTERPRO; IPR000589; -.
DR   PROSITE; PS00362; RIBOSOMAL_S15; PARTIAL.
KW   Ribosomal protein; Chloroplast; rRNA-binding.
FT   NON_TER      20      20
SQ   SEQUENCE   20 AA;  2294 MW;  B613012072A648A3 CRC64;

  Query Match             24.0%;  Score 23;  DB 8;  Length 20;
  Best Local Similarity   44.4%;  Pred. No. 2.9e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 KKKLFLSLL 17
           ||  |:|::
Db       2 KKNSFISVI 10


RESULT  13
O08806
ID   O08806      PRELIMINARY;      PRT;    20 AA.
AC   O08806;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   SERINE PROTEASE INHIBITOR 14 (SERINE PROTEINASE INHIBITOR MNK26)
DE   (FRAGMENT).
GN   SPI14.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Sun J., Ooms L., Bird C.H., Sutton V.R., Trapani J.A., Bird P.I.;
RL   J. Biol. Chem. 0:0-0(0).
DR   EMBL; U96709; AAB57821.1; -.
DR   MGD; MGI:894672; Spi14.
FT   NON_TER      1      1
FT   NON_TER     20     20
SQ   SEQUENCE   20 AA;  2039 MW;  38571040F8263692 CRC64;

  Query Match             24.0%;  Score 23;  DB 11;  Length 20;
  Best Local Similarity   46.7%;  Pred. No. 2.9e+03;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       3 VQASYSKKKLFLSLL 17
           | ||  |  || |::
Db       4 VAASAGKIILFSSMV 18


RESULT  14
P78533
ID   P78533      PRELIMINARY;      PRT;    15 AA.
AC   P78533;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   DEOXYGUANOSINE KINASE (EC 2.7.1.113) (FRAGMENT).
GN   DGK.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Johansson M., Karlsson A.;
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ATP + DEOXYGUANOSINE = ADP + DGMP.
DR   EMBL; U62042; AAB48932.1; -.
KW   Transferase.
FT   NON_TER      15      15
SQ   SEQUENCE   15 AA;  1706 MW;  53575609CC614F8E CRC64;

  Query Match             22.9%;  Score 22;  DB 4;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 3.2e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      11 KLFLSLL 17
           :|||| |
Db       5 RLFLSRL 11


RESULT  15
Q9UD41
ID   Q9UD41      PRELIMINARY;      PRT;    16 AA.
AC   Q9UD41;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   BETA-ISOFORM THYROID HORMONE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95105146.
RA Flynn T.R., Hollenberg A.N., Cohen O., Menke J.B., Usala S.J.;
RA Tollin S., Hegarty M.K., Wondisford F.E.;
RT "A novel C-terminal domain in the thyroid hormone receptor selectively
RT mediates thyroid hormone inhibition.";
RL J. Biol. Chem. 269:32713-32716(1994).
SQ SEQUENCE 16 AA; 1920 MW; 79A831A4F3C8E22F CRC64;

Query Match 22.9%; Score 22; DB 4; Length 16;
Best Local Similarity 40.0%; Pred. No. 3.4e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy      11 KLFLSLLDFQ 20
           || : : | |
Db       4 KLLMKVTDLQ 13
```

Search completed: December 21, 2000, 08:35:40
Job time: 603 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:50 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-36
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	33	30.3	14	1	US-08-311-611A-40	Sequence 40, Appl
2	33	30.3	14	1	US-08-372-783-40	Sequence 40, Appl
3	33	30.3	14	1	US-08-372-105-40	Sequence 40, Appl
4	33	30.3	14	1	US-08-306-473A-40	Sequence 40, Appl
5	33	30.3	14	1	US-08-209-762-40	Sequence 40, Appl
6	33	30.3	14	1	US-08-473-344-40	Sequence 40, Appl
7	33	30.3	14	2	US-08-621-803-39	Sequence 39, Appl
8	33	30.3	14	2	US-08-485-445A-40	Sequence 40, Appl
9	33	30.3	14	2	US-08-621-259A-13	Sequence 13, Appl
10	33	30.3	14	3	US-09-119-263-40	Sequence 40, Appl
11	33	30.3	14	4	PCT-US94-02465-40	Sequence 40, Appl
12	33	30.3	14	4	PCT-US95-00498-40	Sequence 40, Appl
13	33	30.3	14	4	PCT-US95-00656-40	Sequence 40, Appl
14	33	30.3	14	4	PCT-US95-09262-13	Sequence 13, Appl
15	31	28.4	10	2	US-08-621-803-227	Sequence 227, App
16	31	28.4	10	2	US-08-621-259A-219	Sequence 219, App
17	31	28.4	11	2	US-08-621-803-228	Sequence 228, App
18	31	28.4	11	2	US-08-621-259A-220	Sequence 220, App
19	31	28.4	12	2	US-08-621-803-230	Sequence 230, App
20	31	28.4	12	2	US-08-621-259A-222	Sequence 222, App
21	31	28.4	13	1	US-08-261-660A-33	Sequence 33, Appl
22	31	28.4	13	1	US-08-261-660A-42	Sequence 42, Appl
23	31	28.4	13	4	PCT-US94-06931-33	Sequence 33, Appl
24	31	28.4	13	4	PCT-US94-06931-42	Sequence 42, Appl
25	31	28.4	14	1	US-08-311-611A-15	Sequence 15, Appl
26	31	28.4	14	1	US-08-311-611A-33	Sequence 33, Appl
27	31	28.4	14	1	US-08-311-611A-34	Sequence 34, Appl
28	31	28.4	14	1	US-08-311-611A-35	Sequence 35, Appl
29	31	28.4	14	1	US-08-311-611A-36	Sequence 36, Appl
30	31	28.4	14	1	US-08-311-611A-44	Sequence 44, Appl
31	31	28.4	14	1	US-08-311-611A-45	Sequence 45, Appl
32	31	28.4	14	1	US-08-311-611A-46	Sequence 46, Appl
33	31	28.4	14	1	US-08-311-611A-81	Sequence 81, Appl
34	31	28.4	14	1	US-08-311-611A-89	Sequence 89, Appl
35	31	28.4	14	1	US-08-311-611A-91	Sequence 91, Appl
36	31	28.4	14	1	US-08-311-611A-106	Sequence 106, App
37	31	28.4	14	1	US-08-311-611A-107	Sequence 107, App
38	31	28.4	14	1	US-08-311-611A-109	Sequence 109, App
39	31	28.4	14	1	US-08-311-611A-110	Sequence 110, App
40	31	28.4	14	1	US-08-311-611A-122	Sequence 122, App
41	31	28.4	14	1	US-08-311-611A-142	Sequence 142, App
42	31	28.4	14	1	US-08-372-783-15	Sequence 15, Appl
43	31	28.4	14	1	US-08-372-783-33	Sequence 33, Appl
44	31	28.4	14	1	US-08-372-783-34	Sequence 34, Appl
45	31	28.4	14	4	PCT-US95-09262-62	Sequence 62, Appl

ALIGNMENTS

```
RESULT 1
US-08-311-611A-40
; Sequence 40, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
;   APPLICANT: Cohen, Jonathan
;   APPLICANT: Kung, Ada H.C.
;   APPLICANT: Lambert, Jr., Lewis H.
;   TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
;   TITLE OF INVENTION: Infection by Administration of
;   TITLE OF INVENTION: Bactericidal/Permeability-Increasing
;   TITLE OF INVENTION:
;   NUMBER OF SEQUENCES: 227
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
;     STREET: 6300 Sears Tower, 233 South Wacker Drive
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: USA
;     ZIP: 60606-6402
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/311,611A
;     FILING DATE:
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/273,401
;     FILING DATE: 11-JUL-1994
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: 08/125,651
;     FILING DATE: 22-SEP-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Sharp, Jeffrey S.
;     REGISTRATION NUMBER: 31,879
;     REFERENCE/DOCKET NUMBER: 32251
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 312/474-6300
;     TELEFAX: 312/474-0448
;     TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 40:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 14 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FEATURE:
;     NAME/KEY: misc_feature
;     OTHER INFORMATION: "BPI.38"
```

```
US-08-311-611A-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   2
US-08-372-783-40
; Sequence 40, Application US/08372783
; Patent No. 5578572
;  GENERAL INFORMATION:
;    APPLICANT:  Horwitz, Arnold H.
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/372,783
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-08-372-783-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   3
US-08-372-105-40
; Sequence 40, Application US/08372105
; Patent No. 5627153
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Scannon, Patrick J.
;    TITLE OF INVENTION:  Anti-Fungal Materials and Methods
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/372,105
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-08-372-105-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   4
US-08-306-473A-40
; Sequence 40, Application US/08306473A
; Patent No. 5652332
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
```

```
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/
;    TITLE OF INVENTION:  Permeability-Increasing Protein and
;    TITLE OF INVENTION:  Uses Thereof
;    NUMBER OF SEQUENCES:  226
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  Suite 3000, 10 S. Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,473A
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-306-473A-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   5
US-08-209-762-40
; Sequence 40, Application US/08209762
; Patent No. 5733872
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/209,762
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5733872nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-209-762-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   6
US-08-473-344-40
; Sequence 40, Application US/08473344
; Patent No. 5763567
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Banner & Allegretti, Ltd.
;      STREET:  10  South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/473,344
;      FILING DATE:  7-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/306,473
;      FILING DATE:  15-SEP-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/209,762
;      FILING DATE:  11-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McDonnell, John J.
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  93,1133-J
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
```

```
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-473-344-40


  Query Match             30.3%;  Score 33;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   7
US-08-621-803-39
; Sequence 39, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-39


  Query Match             30.3%;  Score 33;  DB 2;  Length 14;
```

```
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   8
US-08-485-445A-40
; Sequence 40, Application US/08485445A
; Patent No. 5856438
;  GENERAL INFORMATION:
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/
;    TITLE OF INVENTION:  Permeability-Increasing Protein and
;    TITLE OF INVENTION:  Uses Thereof
;    NUMBER OF SEQUENCES:  226
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  Suite 3400, 500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,445A
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11018US08/100-224.P4.C1B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-707-8889
;      TELEFAX:  312-707-9155
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
US-08-485-445A-40


  Query Match             30.3%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT   9
US-08-621-259A-13
; Sequence 13, Application US/08621259A
; Patent No. 5858974
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Little II, Roger G
;    APPLICANT:  Lim, Edward
;    APPLICANT:  Fadem, Mitchell B.
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  252
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McAndrews, Held & Malloy, Ltd.
;      STREET:  500 West Madison Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60661
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,259A
;      FILING DATE:  21-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/504,841
;      FILING DATE:  20-JUL-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McNicholas, Janet M.
;      REGISTRATION NUMBER:  32,918
;      REFERENCE/DOCKET NUMBER:  11021US02
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/707-8889
;      TELEFAX:  312/707-9155
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-259A-13


  Query Match             30.3%;  Score 33;  DB 2;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  10
US-09-119-263-40
; Sequence 40, Application US/09119263
; Patent No. 6054431
;  GENERAL INFORMATION:
;    APPLICANT:  Horwitz, Arnold H.
;    APPLICANT:  Lambert, Lewis H.
;    APPLICANT:  Little, Roger G.
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/119,263
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,116
;      FILING DATE:
;      APPLICATION NUMBER:  08/372,783
;      FILING DATE:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
US-09-119-263-40


  Query Match             30.3%;  Score 33;  DB 3;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  11
PCT-US94-02465-40
; Sequence 40, Application PC/TUS9402465
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Biologically Active Peptides from
;    TITLE OF INVENTION:  Functional Domains of Bactericidal/Permeability-Increasing
;    TITLE OF INVENTION:  Protein and Uses Thereof
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/02465
;      FILING DATE:  11-JAN-1994
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Noonan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  93,1133
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "BPI.38"
PCT-US94-02465-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  12
PCT-US95-00498-40
; Sequence 40, Application PC/TUS9500498
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Gram-Positive Bacterial Methods and
;    TITLE OF INVENTION:  Materials
;    NUMBER OF SEQUENCES:  237
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00498
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00498-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  13
PCT-US95-00656-40
; Sequence 40, Application PC/TUS9500656
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Materials and Methods
;    NUMBER OF SEQUENCES:  227
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/00656
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rin-Laures, Li-Hsien
;      REGISTRATION NUMBER:  33,547
;      REFERENCE/DOCKET NUMBER:  27129/32415
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  40:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
PCT-US95-00656-40


  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  14
PCT-US95-09262-13
; Sequence 13, Application PC/TUS9509262
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Anti-Fungal Peptides
;    NUMBER OF SEQUENCES:  206
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/09262
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/372,105
;      FILING DATE:  13-JAN-95
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/306,473
;      FILING DATE:  15-SEP-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/273,540
;      FILING DATE:  11-JUL-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/209,762
;      FILING DATE:  11-MAR-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/183,222
;      FILING DATE:  14-JAN-94
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/093,202
;      FILING DATE:  15-JUL-93
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030,644
;      FILING DATE:  12-MAR-93
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/10040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.38"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated"
PCT-US95-09262-13

  Query Match             30.3%;  Score 33;  DB 4;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 18;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       5 GWLAQLF 11


RESULT  15
US-08-621-803-227
; Sequence 227, Application US/08621803
; Patent No. 5851802
;  GENERAL INFORMATION:
;    APPLICANT:  Better, Marc D.
;    TITLE OF INVENTION:  Methods for Recombinant Microbial Production of
;    TITLE OF INVENTION:  Fusion Proteins and BPI-Derived Peptides
;    NUMBER OF SEQUENCES:  265
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/621,803
;      FILING DATE:  22-MAR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25,447
;      REFERENCE/DOCKET NUMBER:  27129/33199
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  227:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FEATURE:
;      NAME/KEY:  misc_feature
;      OTHER INFORMATION:  "XMP.388"
;    FEATURE:
;      NAME/KEY:  Modified-site
;      LOCATION:  C-Terminus
;      OTHER INFORMATION:  /label= Amidation
;      OTHER INFORMATION:  /note= "The C-Terminus is Amidated."
US-08-621-803-227


  Query Match             28.4%;  Score 31;  DB 2;  Length 10;
  Best Local Similarity   71.4%;  Pred. No. 26;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           ||: |||
Db       1 GWLIQLF 7


Search completed: December 21, 2000, 08:31:50
Job time: 373 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:08 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-36
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	31	28.4	18	2	G02018	proteasome chain L
2	30	27.5	16	2	PH1790	T cell receptor al
3	29	26.6	10	2	S39030	lysyl-bradykinin -
4	27	24.8	15	2	PH1788	T cell receptor al
5	25	22.9	17	2	PH1802	T cell receptor al
6	25	22.9	17	2	PH1822	T cell receptor al
7	25	22.9	18	2	F49215	urease (EC 3.5.1.5
8	24	22.0	4	2	A34626	RPCH-related neuro
9	24	22.0	8	2	A61348	red pigment-concen
10	24	22.0	8	2	A33995	adipokinetic hormo
11	24	22.0	8	2	S55310	adipokinetic hormo
12	24	22.0	8	2	A58620	adipokinetic hormo
13	24	22.0	10	2	A43405	6-phosphofructo-2-
14	24	22.0	10	2	A60421	hypertrehalosemic
15	24	22.0	10	2	S08997	hypertrehalosemic
16	24	22.0	10	2	S08998	hypertrehalosemic
17	24	22.0	10	2	A26381	hypertrehalosemic
18	24	22.0	10	2	B33995	hypotrehalosemic h
19	24	22.0	15	2	A26228	spot 42 protein -
20	24	22.0	15	2	PH1329	Ig heavy chain DJ
21	24	22.0	15	2	S38976	chromogranin A - b
22	24	22.0	16	2	PT0282	Ig heavy chain CDR
23	24	22.0	19	2	S29212	protein C - oat (f
24	24	22.0	20	2	A47687	3-dehydroquinate d
25	23	21.1	7	2	A61081	typtophyllin, basi
26	23	21.1	11	2	S68649	spermadhesin AQN-3
27	23	21.1	15	2	S67975	apolipoprotein Cb2
28	23	21.1	16	1	A49761	locustapyrokinin -
29	23	21.1	17	2	PH0778	T-cell receptor al
30	23	21.1	18	2	S71592	serine proteinase
31	23	21.1	19	2	PH1339	Ig heavy chain DJ
32	22	20.2	9	2	S78426	52.5K protein - sp
33	22	20.2	9	2	G56978	collagen alpha 1(I
34	22	20.2	11	2	D56979	collagen alpha 1(I
35	22	20.2	11	2	PD0442	NIPSNAP2 protein -
36	22	20.2	12	2	PQ0776	NADH dehydrogenase
37	22	20.2	13	2	A28505	neurotensin-like p
38	22	20.2	14	2	PH1347	Ig heavy chain DJ
39	22	20.2	14	2	PH1311	Ig heavy chain DJ
40	22	20.2	14	2	PH1321	Ig heavy chain DJ
41	22	20.2	14	2	PH1305	Ig heavy chain DJ
42	22	20.2	14	2	PH1306	Ig heavy chain DJ
43	22	20.2	15	2	PT0222	Ig heavy chain CDR
44	22	20.2	15	2	PH1314	Ig heavy chain DJ
45	22	20.2	15	2	PH1310	Ig heavy chain DJ

ALIGNMENTS

RESULT 1
G02018
proteasome chain LMP7 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 22-Jun-1999
C;Accession: G02018
R;Kim, T.
submitted to the EMBL Data Library, July 1995
A;Reference number: G09054
A;Accession: G02018
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <KIM>
A;Cross-references: EMBL:U32862; NID:g1045468; PIDN:AAA80234.1; PID:g1045469
C;Genetics:
A;Gene: LMP7
C;Superfamily: multicatalytic endopeptidase complex chain C9

Query Match 28.4%; Score 31; DB 2; Length 18;
Best Local Similarity 57.1%; Pred. No. 1e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Qy       9 REPGWIK 15
            :| ||:|
Db       4 KEDGWVK 10
```

RESULT 2
PH1790
T cell receptor alpha chain V region (clone 2PBL V alpha 24-6) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1790
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1790
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-16 <POR>

Query Match 27.5%; Score 30; DB 2; Length 16;
Best Local Similarity 75.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 HLQGEREP 11
            || ||| |
Db       1 HLCGERGP 8
```

```
RESULT   3
S39030
lysyl-bradykinin - rainbow trout
C;Species: Oncorhynchus mykiss (rainbow trout)
C;Date: 19-May-1994 #sequence_revision 19-Apr-1996 #text_change 19-Apr-1996
C;Accession: S39030
R;Conlon, J.M.; Olson, K.R.
FEBS Lett. 334, 75-78, 1993
A;Title: Purification of a vasoactive peptide related to lysyl-bradykinin from trout pla
A;Reference number: S39030; MUID:94039817
A;Accession: S39030
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <CON>

  Query Match             26.6%;  Score 29;  DB 2;  Length 10;
  Best Local Similarity   66.7%;  Pred. No. 1.2e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 EREPGW 13
           :| |||
Db       1 KRPPGW 6


RESULT   4
PH1788
T cell receptor alpha chain V region (clone 2PBL V alpha 24-4) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1788
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1788
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-15 <POR>

  Query Match             24.8%;  Score 27;  DB 2;  Length 15;
  Best Local Similarity   71.4%;  Pred. No. 3.7e+02;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGERE 10
           || |||:
Db       1 HLCGERQ 7


RESULT   5
PH1802
T cell receptor alpha chain V region (clone 3PBL V alpha 24-8) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1802
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blood
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1802
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>

  Query Match             22.9%;  Score 25;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 8.8e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGER 9
           || |||
Db       1 HLCGER 6


RESULT   6
PH1822
T cell receptor alpha chain V region (clone 5PBL V alpha 24-5) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999
C;Accession: PH1822
R;Porcelli, S.; Yockey, C.E.; Brenner, M.B.; Balk, S.P.
J. Exp. Med. 178, 1-16, 1993
A;Title: Analysis of T cell antigen receptor (TCR) expression by human peripheral blo
.
A;Reference number: PH1754; MUID:93301585
A;Accession: PH1822
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-17 <POR>

  Query Match             22.9%;  Score 25;  DB 2;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 8.8e+02;
  Matches    5;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 HLQGER 9
           || |||
Db       1 HLCGER 6


RESULT   7
F49215
urease (EC 3.5.1.5) small chain UreA - Helicobacter mustelae (ATCC 43772) (fragment)
C;Species: Helicobacter mustelae
C;Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 30-May-1997
C;Accession: F49215
R;Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A;Title: Purification and characterization of the urease enzymes of Helicobacter spec
A;Reference number: A49215; MUID:93084378
A;Accession: F49215
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-18 <TUR>
A;Note: sequence extracted from NCBI backbone (NCBIP:119487)
C;Superfamily: urease 26K chain; urease 11K chain homology; urease 12K chain homology
C;Keywords: hydrolase

  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 9.4e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGE 8
           ::||  ||
Db       9 MMLHYAGE 16


RESULT   8
A34626
RPCH-related neuropeptide - ferruginous spindle
C;Species: Fusinus ferrugineus (ferruginous spindle)
C;Date: 06-Jul-1990 #sequence_revision 06-Jul-1990 #text_change 31-Dec-1993
C;Accession: A34626
R;Kuroki, Y.; Kanda, T.; Kubota, I.; Fujisawa, Y.; Ikeda, T.; Miura, A.; Minamitake,
Biochem. Biophys. Res. Commun. 167, 273-279, 1990
A;Title: A molluscan neuropeptide related to the crustacean hormone, RPCH.
A;Reference number: A34626; MUID:90179762
A;Accession: A34626
A;Status: preliminary
A;Molecule type: protein
```

A;Residues: 1-4 <KUR>
C;Keywords: neuropeptide

Query Match 22.0%; Score 24; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 2 PGW 4

RESULT 9
A61348
red pigment-concentrating hormone - northern shrimp
N;Alternate names: blanching hormone
C;Species: Pandalus borealis (northern shrimp)
C;Date: 02-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 07-May-1999
C;Accession: A61348; S07139
R;Fernlund, P.; Josefsson, L.
Science 177, 173-175, 1972
A;Title: Crustacean color-change hormone: amino acid sequence and chemical synthesis.
A;Reference number: A61348; MUID:72228738
A;Accession: A61348
A;Molecule type: protein
A;Residues: 1-8 <FER1>
R;Fernlund, P.
Biochim. Biophys. Acta 371, 304-311, 1974
A;Title: Structure of the red-pigment-concentrating hormone of the shrimp, Pandalus bore
A;Reference number: S07139; MUID:75054965
A;Accession: S07139
A;Molecule type: protein
A;Residues: 'E',2-8 <FER2>
A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
C;Comment: This neuropeptide, isolated from the eyestalks of the shrimp and active in pi
zed pigment-containing cells.
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; blocked amino end; hormone; neuropeptide; pyroglutami
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 10
A33995
adipokinetic hormone - black horse fly
C;Species: Tabanus atratus (black horse fly)
C;Date: 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change 31-Oct-1997
C;Accession: A33995
R;Jaffe, H.; Raina, A.K.; Riley, C.T.; Fraser, B.A.; Nachman, R.J.; Vogel, V.W.; Zhang,
Proc. Natl. Acad. Sci. U.S.A. 86, 8161-8164, 1989
A;Title: Primary structure of two neuropeptide hormones with adipokinetic and hypotrehal
A;Reference number: A33995; MUID:90046758
A;Accession: A33995
A;Molecule type: protein
A;Residues: 1-8 <JAF>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; corpora cardiaca; hormone; neuropeptide; pyroglutamic
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status predicted
F;8/Modified site: amidated carboxyl end (Trp) #status predicted

Query Match 22.0%; Score 24; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 11
S55310
adipokinetic hormone - damselfly (Pseudagrion inconspicuum)
N;Alternate names: Psi-AKH
C;Species: Pseudagrion inconspicuum
C;Date: 19-Mar-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: S55310
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
A;Accession: S55310
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 12
A58620
adipokinetic hormone - damselfly (Ischnura senegalensis)
C;Species: Ischnura senegalensis
C;Date: 28-Oct-1997 #sequence_revision 31-Oct-1997 #text_change 07-May-1999
C;Accession: A58620
R;Janssens, M.P.E.; Kellner, R.; Gaede, G.
Biochem. J. 302, 539-543, 1994
A;Title: A novel adipokinetic octapeptide found in the damselflies Pseudagrion incons
A;Reference number: S55310; MUID:94379987
A;Accession: A58620
A;Molecule type: protein
A;Residues: 1-8 <JAN>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;8/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 13
A43405
6-phosphofructo-2-kinase (EC 2.7.1.105) / fructose-2,6-bisphosphate 2-phosphatase (EC
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 31-Mar-2000
C;Accession: A43405
R;Ventura, F.; Rosa, J.L.; Ambrosio, S.; Pilkis, S.J.; Bartrons, R.
J. Biol. Chem. 267, 17939-17943, 1992
A;Title: Bovine brain 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase. Evidence

A;Reference number: A43405; MUID:92388154
A;Accession: A43405
A;Molecule type: protein
A;Residues: 1-10 <VEN>
C;Superfamily: 6-phosphofructo-2-kinase / fructose-2,6-bisphosphate 2-phosphatase; phosp
C;Keywords: phosphoric monoester hydrolase; phosphotransferase

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 80.0%; Pred. No. 7.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 HLQGE 8
||:||
Db 5 HLKGE 9

RESULT 14
A60421
hypertrehalosemic hormone - German cockroach
N;Alternate names: Bld-HrTH
C;Species: Blattella germanica (German cockroach)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 31-Oct-1997
C;Accession: A60421; S09137
R;Veenstra, J.A.; Camps, F.
Neuropeptides 15, 107-109, 1990
A;Title: Structure of the hypertrehalosemic neuropeptide of the German cockroach, Blatte
A;Reference number: A60421; MUID:91179584
A;Accession: A60421
A;Molecule type: protein
A;Residues: 1-10 <VEE>
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659
A;Accession: S09137
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

RESULT 15
S08997
hypertrehalosemic neuropeptide Bld-HrTH - cockroach (Gromphadorina portentosa)
C;Species: Gromphadorina portentosa
C;Date: 30-Jun-1992 #sequence_revision 14-Sep-1994 #text_change 24-Oct-1997
C;Accession: S08997
R;Gaede, G.; Rinehart, K.L.
Biol. Chem. Hoppe-Seyler 371, 345-354, 1990
A;Title: Primary structures of hypertrehalosaemic neuropeptides isolated from the corpor
entalis and of the stick insect Extatosoma tiaratum assigned by tandem fast atom bombard
A;Reference number: S08995; MUID:90253659
A;Accession: S08997
A;Molecule type: protein
A;Residues: 1-10 <GAE>
C;Superfamily: adipokinetic hormone
C;Keywords: amidated carboxyl end; hormone; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;10/Modified site: amidated carboxyl end (Thr) #status experimental

Query Match 22.0%; Score 24; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 7.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 PGW 13
|||
Db 6 PGW 8

Search completed: December 21, 2000, 08:30:08
Job time: 272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:33:00 ; Search time 52.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-36
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	25.7	19	1	TRP3_LEUMA	P81735 leucophaea
2	24	22.0	8	1	AKH_TABAT	P14595 tabanus atr
3	24	22.0	8	1	RPCH_PANBO	P08939 pandalus bo
4	24	22.0	10	1	HTF_NAUCI	P10939 nauphoeta c
5	24	22.0	10	1	HTF_TABAT	P14596 tabanus atr
6	24	22.0	20	1	AROQ_AMYME	P46380 amycolatops
7	23	21.1	16	1	LPK1_LOCMI	P20404 locusta mig
8	22	20.2	13	1	BRK_PARID	P42717 parapolybia
9	22	20.2	13	1	NEUT_CHICK	P13724 gallus gall
10	21	19.3	10	1	Q2OG_COMTE	P80466 comamonas t
11	21	19.3	13	1	CXET_CONTE	P81755 conus texti
12	21	19.3	17	1	CHH3_BOMMO	P20729 bombyx mori
13	21	19.3	20	1	CRTC_SPIOL	P30806 spinacia ol
14	20	18.3	12	1	PORD_METTM	P80903 methanobact
15	20	18.3	13	1	HPA1_RANES	P32415 rana escule
16	20	18.3	14	1	CAT2_FASHE	P80342 fasciola he
17	20	18.3	15	1	TA1_TREBR	P34070 tremella br
18	20	18.3	19	1	CAT3_FASHE	P80532 fasciola he
19	20	18.3	19	1	LANA_ACTLG	P56650 actinoplane
20	20	18.3	20	1	CAT1_FASHE	Q09093 fasciola he
21	19.5	17.9	20	1	COG1_PARCM	P20731 paralithode
22	19	17.4	7	1	UH11_RAT	P56576 rattus norv
23	19	17.4	10	1	GON1_PETMA	P04378 petromyzon
24	19	17.4	10	1	GON3_ONCKE	P20367 oncorhynchu
25	19	17.4	10	1	GONL_SQUAC	P27429 squalus aca
26	19	17.4	12	1	RS19_ELYEP	Q47881 elm yellows
27	19	17.4	13	1	FARB_ASCSU	P43173 ascaris suu
28	19	17.4	13	1	NEUT_RANTE	P41536 rana tempor
29	19	17.4	14	1	RS19_CLOPP	Q46228 clover prol
30	19	17.4	14	1	RS19_LOWBP	Q48878 loofah witc
31	19	17.4	15	1	VORA_METTM	P80907 methanobact
32	19	17.4	18	1	NPA_BOVIN	P15506 bos taurus
33	19	17.4	19	1	FIBB_HORSE	P14471 equus cabal
34	19	17.4	19	1	PYRB_PSEFL	P56585 pseudomonas
35	19	17.4	20	1	COG1_CHIOP	P34153 chionoecete
36	18	16.5	8	1	AKH_LIBAU	P25418 libellula a
37	18	16.5	8	1	CCKN_MACEU	P30369 macropus eu
38	18	16.5	8	1	HTF1_PERAM	P04548 periplaneta
39	18	16.5	8	1	HTF2_PERAM	P04549 periplaneta
40	18	16.5	8	1	HTF_TENMO	P25419 tenebrio mo
41	18	16.5	8	1	UF06_MOUSE	P38644 mus musculu
42	18	16.5	10	1	CAER_LITXA	P56264 litoria xan
43	18	16.5	10	1	HTF1_ROMMI	P18110 romalea mic
44	18	16.5	10	1	HTF2_CARMO	P11385 carausius m
45	18	16.5	11	1	BRK_MEGFL	P12797 megascolia

ALIGNMENTS

RESULT 1
TRP3_LEUMA
ID TRP3_LEUMA STANDARD; PRT; 19 AA.
AC P81735;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE TACHYKININ-RELATED PEPTIDE 3 (LEMTRP 3).
OS Leucophaea maderae (Madeira cockroach).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC Blaberoidea; Blaberidae; Leucophaea.
RN [1]
RP SEQUENCE.
RC TISSUE=MIDGUT;
RX MEDLINE; 97053012.
RA Muren J.E., Naessel D.R.;
RT "Isolation of five tachykinin-related peptides from the midgut of
RT the cockroach Leucophaea madera: existence of N-terminally extended
RT isoforms.";
RL Regul. Pept. 65:185-196(1996).
CC -!- FUNCTION: MYOACTIVE PEPTIDE. INCREASES THE AMPLITUDE AND FREQUENCY
CC OF SPONTANEOUS CONTRACTIONS AND TONUS OF HINDGUT MUSCLE.
CC -!- TISSUE SPECIFICITY: MIDGUT.
CC -!- SIMILARITY: SOME SIMILARITY TO TACHYKININS.
KW Tachykinin; Neuropeptide; Amidation.
FT MOD_RES 19 19 AMIDATION.
SQ SEQUENCE 19 AA; 1930 MW; 99B5471A011625E5 CRC64;

Query Match 25.7%; Score 28; DB 1; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      7 GEREPG 12
          ||| ||
Db      2 GERAPG 7
```

RESULT 2
AKH_TABAT
ID AKH_TABAT STANDARD; PRT; 8 AA.
AC P14595;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE ADIPOKINETIC HORMONE (AKH) (DIPTERAN CORPORA CARDIACA FACTOR I)
DE (DCC I).
OS Tabanus atratus (Horse fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC Tabanidae; Tabanus.
RN [1]
RP SEQUENCE.
RC TISSUE=CORPORA CARDIACA;
RX MEDLINE; 90046758.

```
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC       CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC       DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC       MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A33995; A33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation; Flight.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  949 MW;  86786771A9D1A736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   3
RPCH_PANBO
ID   RPCH_PANBO     STANDARD;      PRT;     8 AA.
AC   P08939;
DT   01-NOV-1988 (Rel. 09, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   RED PIGMENT CONCENTRATING HORMONE (RPCH).
OS   Pandalus borealis (Northern red shrimp).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidea;
OC   Pandalidae; Pandalus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 75054965.
RA   Fernlund P.;
RT   "Structure of the red-pigment-concentrating hormone of the shrimp,
RT   Pandalus borealis.";
RL   Biochim. Biophys. Acta 371:304-311(1974).
CC   -!- FUNCTION: THIS HORMONE ADAPTS THE ANIMAL TO LIGHT BACKGROUNDS BY
CC       STIMULATING CONCENTRATION OF THE PIGMENT OF ITS RED BODY-
CC       CHROMATOPHORES.
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; S07139; S07139.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Pigment; Hormone; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  948 MW;  86786775B9C44736 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   4
HTF_NAUCI
ID   HTF_NAUCI      STANDARD;      PRT;     10 AA.
AC   P10939;
DT   01-JUL-1989 (Rel. 11, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   HYPERTREHALOSAEMIC HORMONE (HTH) (HYPERTREHALOSAEMIC NEUROPEPTIDE).
OS   Nauphoeta cinerea (Cinereous cockroach) (Gray cockroach),
OS   Leucophaea maderae (Madeira cockroach),
OS   Blattella germanica (German cockroach), and
OS   Gromphadorina portentosa (Cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blaberoidea; Blaberidae; Nauphoeta.
RN   [1]
RP   SEQUENCE.
RC   SPECIES=N.CINEREA; TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 87100208.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Amino acid sequence of a hypertrehalosaemic neuropeptide from the
RT   corpus cardiacum of the cockroach, Nauphoeta cinerea.";
RL   Biochem. Biophys. Res. Commun. 141:774-781(1986).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=L.MADERAE, G.PORTENTOSA, AND B.GERMANICA;
RX   MEDLINE; 90253659.
RA   Gaede G., Rinehart K.L. Jr.;
RT   "Primary structures of hypertrehalosaemic neuropeptides isolated from
RT   the corpora cardiaca of the cockroaches Leucophaea maderae,
RT   Gromphadorhina portentosa, Blattella germanica and Blatta orientalis
RT   and of the stick insect Extatosoma tiaratum assigned by tandem fast
RT   atom bombardment mass spectrometry.";
RL   Biol. Chem. Hoppe-Seyler 371:345-354(1990).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=B.GERMANICA;
RX   MEDLINE; 91179584.
RA   Veenstra J.A., Camps F.;
RT   "Structure of the hypertrehalosemic neuropeptide of the German
RT   cockroach, Blattella germanica.";
RL   Neuropeptides 15:107-109(1990).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; A26381; A26381.
DR   PIR; S08997; S08997.
DR   PIR; S08998; S08998.
DR   PIR; S09137; S09137.
DR   PIR; A60421; A60421.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1092 MW;  056236786775B9C4 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       6 PGW 8


RESULT   5
HTF_TABAT
ID   HTF_TABAT      STANDARD;      PRT;     10 AA.
AC   P14596;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   HYPERTREHALOSAEMIC FACTOR (HOTH) (DIPTERAN CORPORA CARDIACA FACTOR II)
```

```
DE   (DCC II).
OS   Tabanus atratus (Horse fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Tabanomorpha;
OC   Tabanidae; Tabanus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 90046758.
RA   Jaffe H., Raina A.K., Riley C.T., Fraser B.A., Nachman R.J.,
RA   Vogel V.W., Zhang Y.-S., Hayes D.K.;
RT   "Primary structure of two neuropeptide hormones with adipokinetic and
RT   hypotrehalosemic activity isolated from the corpora cardiaca of horse
RT   flies (Diptera).";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:8161-8164(1989).
CC   -!- FUNCTION: HYPERTREHALOSAEMIC FACTORS ARE NEUROPEPTIDES THAT
CC       ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
CC       THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
CC   -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR   PIR; B33995; B33995.
DR   INTERPRO; IPR002047; -.
DR   PROSITE; PS00256; AKH; 1.
KW   Neuropeptide; Amidation.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1169 MW;  916036786771A9D1 CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 10;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+02;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     11 PGW 13
          |||
Db      6 PGW 8


RESULT   6
AROQ_AMYME
ID   AROQ_AMYME     STANDARD;      PRT;     20 AA.
AC   P46380;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10) (3-DEHYDROQUINASE)
DE   (TYPE II DHQASE) (FRAGMENT).
GN   AROQ.
OS   Amycolatopsis methanolica.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC   Amycolatopsis.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=NCIB 11946;
RX   MEDLINE; 93123995.
RA   Euverink G.J.W., Hessels G.I., Vrijbloed J.W., Coggins J.R.,
RA   Dijkhuizen L.;
RT   "Purification and characterization of a dual function
RT   3-dehydroquinate dehydratase from Amycolatopsis methanolica.";
RL   J. Gen. Microbiol. 138:2449-2457(1992).
CC   -!- FUNCTION: CATALYZE A TRANS-DEHYDRATION VIA AN ENOLATE
CC       INTERMEDIATE. IS INVOLVED IN BOTH THE CATABOLISM OF QUINATE AND
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS. HAS A TEMPERATURE
CC       OPTIMAL OF 76 DEGREES CELSIUS AND PH OPTIMAL OF 9.0. TYPE II
CC       ENZYMES ARE THERMOSTABLE.
CC   -!- CATALYTIC ACTIVITY: 3-DEHYDROQUINATE = 3-DEHYDROSHIKIMATE + H(2)O.
CC   -!- PATHWAY: THIRD STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN
CC       THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).
CC   -!- SUBUNIT: HOMODODECAMER.
CC   -!- SIMILARITY: BELONGS TO THE TYPE-II 3-DEHYDROQUINASE FAMILY.
DR   INTERPRO; IPR001874; -.
DR   PROSITE; PS01029; DEHYDROQUINASE_II; PARTIAL.
KW   Quinate metabolism; Aromatic amino acid biosynthesis; Lyase.
FT   NON_TER       20     20
SQ   SEQUENCE   20 AA;  2197 MW;  C24AA183E5CFDF0A CRC64;


  Query Match             22.0%;  Score 24;  DB 1;  Length 20;
  Best Local Similarity   80.0%;  Pred. No. 7e+02;
  Matches    4;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 GEREP 11
          |:|||
Db     15 GKREP 19


RESULT   7
LPK1_LOCMI
ID   LPK1_LOCMI     STANDARD;      PRT;     16 AA.
AC   P20404;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   LOCUSTAPYROKININ 1 (LOM-PK-1).
OS   Locusta migratoria (Migratory locust).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
OC   Acridomorpha; Acridoidea; Acrididae; Locusta.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CORPORA CARDIACA;
RX   MEDLINE; 91224474.
RA   Schoofs L., Holman G.M., Hayes T.K., Nachman R.J., de Loof A.;
RT   "Isolation, primary structure, and synthesis of locustapyrokinin: a
RT   myotropic peptide of Locusta migratoria.";
RL   Gen. Comp. Endocrinol. 81:97-104(1991).
CC   -!- FUNCTION: MEDIATES VISCERAL MUSCLE CONTRACTILE ACTIVITY
CC       (MYOTROPIC ACTIVITY).
CC   -!- SIMILARITY: BELONGS TO THE PYROKININ FAMILY.
DR   PIR; A49761; A49761.
DR   INTERPRO; IPR001484; -.
DR   PROSITE; PS00539; PYROKININ; 1.
KW   Neuropeptide; Amidation; Pyrokinin.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      16     16       AMIDATION.
SQ   SEQUENCE   16 AA;  1827 MW;  A7178BBDCA0AFDD6 CRC64;


  Query Match             21.1%;  Score 23;  DB 1;  Length 16;
  Best Local Similarity   57.1%;  Pred. No. 8e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GWIKQLF 18
          || :| |
Db      6 GWPQQPF 12


RESULT   8
BRK_PARID
ID   BRK_PARID      STANDARD;      PRT;     13 AA.
AC   P42717;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   WASPKININ.
OS   Parapolybia indica.
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Polistinae; Parapolybia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Toki T., Yasuhara T., Nakajima T.;
RT   "Isolation and sequential analysis of peptides on the venom sac of
RT   Parapolybia indica.";
RL   Eisei Dobutsu 39:105-111(1988).
```

```
CC   -!- FUNCTION: SMOOTH MUSCLE CONTRACTION.
KW   Bradykinin; Vasodilator.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1573 MW;  2673CB3D83ECC867 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   50.0%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 QGEREPGW 13
           | :| ||:
Db       1 QZKRPPGF 8


RESULT   9
NEUT_CHICK
ID   NEUT_CHICK     STANDARD;      PRT;     13 AA.
AC   P13724;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   NEUROTENSIN (NT).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 88063566.
RA   Iwabuchi H., Komori S., Ohashi H., Kimura S.;
RT   "The amino acid sequence of a smooth muscle-contracting peptide from
RT   chicken rectum. Identity to chicken neurotensin.";
RL   Jpn. J. Pharmacol. 44:455-459(1987).
CC   -!- FUNCTION: SMOOTH MUSCLE-CONTRACTING PEPTIDE.
DR   PIR; A28505; A28505.
KW   Vasoactive.
FT   MOD_RES       1      1       PYRROLIDONE CARBOXYLIC ACID.
SQ   SEQUENCE   13 AA;  1608 MW;  4C949E714C410DD3 CRC64;


  Query Match             20.2%;  Score 22;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 9.2e+02;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 LHLQGEREP 11
           ||:   | |
Db       2 LHVNKARRP 10


RESULT  10
Q2OG_COMTE
ID   Q2OG_COMTE     STANDARD;      PRT;     10 AA.
AC   P80466;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   QUINOLINE 2-OXIDOREDUCTASE, GAMMA CHAIN (EC 1.-.-.-) (FRAGMENT).
OS   Comamonas testosteroni (Pseudomonas testosteroni).
OC   Bacteria; Proteobacteria; beta subdivision; Comamonadaceae; Comamonas.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=63;
RX   MEDLINE; 96035889.
RA   Schach S., Tshisuaka B., Fetzner S., Lingens F.;
RT   "Quinoline 2-oxidoreductase and 2-oxo-1,2-dihydroquinoline 5,6-
RT   dioxygenase from Comamonas testosteroni 63. The first two enzymes in
RT   quinoline and 3-methylquinoline degradation.";
RL   Eur. J. Biochem. 232:536-544(1995).
CC   -!- FUNCTION: CONVERTS (3-METHYL-)-QUINOLINE TO (3-METHYL-)2-OXO-
CC       1,2-DIHYDROQUINOLINE.
CC   -!- COFACTOR: FAD AND MOLYBDENUM.
CC   -!- PATHWAY: FIRST STEP IN THE DEGRADATION OF QUINOLINE AND
CC       (3-METHYL-)QUINOLINE.
CC   -!- SUBUNIT: HETEROHEXAMER OF TWO ALPHA CHAINS, TWO BETA CHAINS, AND
CC       TWO GAMMA CHAINS (PROBABLE).
KW   Oxidoreductase; Flavoprotein; FAD; Molybdenum.
FT   NON_TER       10     10
SQ   SEQUENCE   10 AA;  1153 MW;  C848CE64433B1DC6 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 10;
  Best Local Similarity   42.9%;  Pred. No. 1e+03;
  Matches    3;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 LQGEREP 11
           :| |: |
Db       2 IQAEKNP 8


RESULT  11
CXET_CONTE
ID   CXET_CONTE     STANDARD;      PRT;     13 AA.
AC   P81755;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   01-OCT-2000 (Rel. 40, Last annotation update)
DE   EPSILON-CONOTOXIN TXIX.
OS   Conus textile (Cloth-of-gold cone).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
OC   Neogastropoda; Conoidea; Conidae; Conus.
RN   [1]
RP   SEQUENCE, CARBOHYDRATE-LINKAGE SITE THR-10, AND STRUCTURE BY NMR.
RC   TISSUE=VENOM;
RX   MEDLINE; 99254114.
RA   Rigby A.C., Lucas-Meunier E., Kalume D.E., Czerwiec E., Hambe B.,
RA   Dahlqvist I., Fossier P., Baux G., Roepstorff P., Baleja J.D.,
RA   Furie B.C., Furie B., Stenflo J.P.;
RT   "A conotoxin from Conus textile with unusual posttranslational
RT   modifications reduces presynaptic Ca2+ influx.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:5758-5763(1999).
CC   -!- FUNCTION: CONOTOXIN WHICH ACTS AT PRESYNAPTIC MEMBRANES, BLOCKING
CC       THE CALCIUM CHANNELS.
CC   -!- PTM: O-GLYCAN CONSISTS OF THE DISACCHARIDE GAL-GALNAC.
DR   PDB; 1WCT; 08-JUN-99.
KW   Presynaptic neurotoxin; Calcium channel inhibitor; Venom; Vitamin K;
KW   Gamma-carboxyglutamic acid; Glycoprotein; Hydroxylation; Bromination;
KW   3D-structure.
FT   DISULFID      2      8
FT   DISULFID      3      9
FT   MOD_RES       1      1       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       4      4       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES       7      7       BROMINATION.
FT   MOD_RES      13     13       HYDROXYLATION.
FT   CARBOHYD     10     10       O-LINKED (GALNAC. . .).
SQ   SEQUENCE   13 AA;  1388 MW;  386C9E1C74AFA378 CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   75.0%;  Pred. No. 1.3e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 EPGW 13
           | ||
Db       4 EDGW 7


RESULT  12
CHH3_BOMMO
ID   CHH3_BOMMO     STANDARD;      PRT;     17 AA.
AC   P20729;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   CHORION CLASS HIGH-CYSTEINE HCA PROTEIN 13 PRECURSOR (HC-A.13)
DE   (FRAGMENT).
```

```
OS   Bombyx mori (Silk moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Bombycoidea; Bombycidae; Bombyx.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 85083111.
RA   Rodakis G.C., Lecanidou R., Eickbush T.H.;
RT   "Diversity in a chorion multigene family created by tandem
RT   duplications and a putative gene-conversion event.";
RL   J. Mol. Evol. 20:265-273(1984).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC       SILK MOTH.
CC   -!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
CC       BELONG CLASSES A, CA AND HCA.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01068; CAB57790.1; -.
DR   PIR; B23219; B23219.
KW   Eggshell; Chorion; Repeat; Multigene family; Signal.
FT   SIGNAL       1    >17
FT   NON_TER     17     17
SQ   SEQUENCE   17 AA;  1913 MW;  5E634508C5355C9C CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 17;
  Best Local Similarity   71.4%;  Pred. No. 1.7e+03;
  Matches    5;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 LLLHLQG 7
           ||| :||
Db       7 LLLCVQG 13


RESULT  13
CRTC_SPIOL
ID   CRTC_SPIOL     STANDARD;      PRT;    20 AA.
AC   P30806;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUL-1993 (Rel. 26, Last annotation update)
DE   CALRETICULIN-LIKE PROTEIN (FRAGMENT).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae;
OC   Caryophyllales; Chenopodiaceae; Spinacia.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=LEAF;
RX   MEDLINE; 93176159.
RA   Menegazzi P., Guzzo F., Baldan B., Mariani P., Treves S.;
RT   "Purification of calreticulin-like protein(s) from spinach leaves.";
RL   Biochem. Biophys. Res. Commun. 190:1130-1135(1993).
CC   -!- FUNCTION: THIS PROTEIN BINDS CALCIUM.
CC   -!- PTM: GLYCOSYLATED.
CC   -!- SIMILARITY: BELONGS TO THE CALRETICULIN FAMILY.
DR   PIR; PC1241; PC1241.
DR   PIR; PC1240; PC1240.
DR   INTERPRO; IPR001580; -.
DR   PROSITE; PS00803; CALRETICULIN_1; PARTIAL.
DR   PROSITE; PS00804; CALRETICULIN_2; PARTIAL.
DR   PROSITE; PS00805; CALRETICULIN_REPEAT; PARTIAL.
KW   Endoplasmic reticulum; Calcium-binding; Glycoprotein.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2645 MW;  00FAB4C9DEEDCB0F CRC64;


  Query Match             19.3%;  Score 21;  DB 1;  Length 20;
  Best Local Similarity   75.0%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      10 EPGW 13
             | ||
Db       9 EDGW 12


RESULT  14
PORD_METTM
ID   PORD_METTM     STANDARD;      PRT;    12 AA.
AC   P80903;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   PYRUVATE SYNTHASE SUBUNIT PORD (EC 1.2.7.1) (PYRUVATE OXIDOREDUCTASE
DE   DELTA CHAIN) (POR) (PYRUVIC-FERREDOXIN OXIDOREDUCTASE DELTA SUBUNIT)
DE   (FRAGMENT).
GN   PORD.
OS   Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC   Methanobacterium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 97261844.
RA   Tersteegen A., Linder D., Thauer R.K., Hedderich R.;
RT   "Structures and functions of four anabolic 2-oxoacid oxidoreductases
RT   in Methanobacterium thermoautotrophicum.";
RL   Eur. J. Biochem. 244:862-868(1997).
CC   -!- FUNCTION: THE PH OPTIMUM IS PH 10.0 AND THE OPTIMAL TEMPERATURE
CC       IS 80 DEGREES CELSIUS.
CC   -!- CATALYTIC ACTIVITY: PYRUVATE + COA + OXIDIZED FERREDOXIN =
CC       ACETYL-COA + CO(2) + REDUCED FERREDOXIN.
CC   -!- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (BY SIMILARITY).
CC   -!- SUBUNIT: HETEROTETRAMER OF ONE ALPHA, ONE BETA, ONE DELTA AND ONE
CC       GAMMA CHAIN.
CC   -!- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.
KW   Oxidoreductase; Electron transport; Iron-sulfur; Repeat; 4Fe-4S.
FT   NON_TER      12     12
SQ   SEQUENCE   12 AA;  1241 MW;  2D54065D1BD1ADD8 CRC64;


  Query Match             18.3%;  Score 20;  DB 1;  Length 12;
  Best Local Similarity   75.0%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       9 REPG 12
           :|||
Db       9 KEPG 12


RESULT  15
HPA1_RANES
ID   HPA1_RANES     STANDARD;      PRT;    13 AA.
AC   P32415;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-OCT-1993 (Rel. 27, Last sequence update)
DT   01-OCT-1993 (Rel. 27, Last annotation update)
DE   HEMOLYTIC PROTEIN A1 (FRAGMENT).
OS   Rana esculenta (Edible frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN SECRETION;
RX   MEDLINE; 90198965.
RA   Simmaco M., de Biase D., Severini C., Aita M., Erspamer G.F.,
RA   Barra D., Bossa F.;
RT   "Purification and characterization of bioactive peptides from skin
RT   extracts of Rana esculenta.";
RL   Biochim. Biophys. Acta 1033:318-323(1990).
```

```
CC   -!- FUNCTION: SHOWS HEMOLYTIC ACTIVITY.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- TISSUE SPECIFICITY: SKIN.
DR   PIR; S09018; S09018.
KW   Amphibian skin; Amidation; Hemolysis.
FT   MOD_RES      13     13       AMIDATION.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1390 MW;  C6BA768B9DFE587D CRC64;


  Query Match             18.3%;  Score 20;  DB 1;  Length 13;
  Best Local Similarity   57.1%;  Pred. No. 1.9e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GWIKQLF 18
           | : |||
Db       7 GILSQLF 13


Search completed: December 21, 2000, 08:33:00
Job time: 443 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:40 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-36
Perfect score: 109
Sequence: 1 LLLHLQGEREPGWIKQLFTN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	26.6	10	13	Q9PRZ1	Q9prz1 oncorhynchu
2	29	26.6	20	4	Q9UC71	Q9uc71 homo sapien
3	28	25.7	9	13	Q9PRJ4	Q9prj4 lepisosteus
4	28	25.7	14	5	O18502	O18502 schistosoma
5	27	24.8	20	2	Q9R4W5	Q9r4w5 helicobacte
6	26	23.9	11	7	O77893	O77893 oreochromis
7	26	23.9	17	2	Q9X3I3	Q9x3i3 prochloroco
8	26	23.9	17	8	Q9T2R9	Q9t2r9 solanum tub
9	25.5	23.4	19	4	Q9UCG3	Q9ucg3 homo sapien
10	25	22.9	14	4	Q16045	Q16045 homo sapien
11	25	22.9	18	2	Q9R5F6	Q9r5f6 helicobacte
12	25	22.9	18	13	Q9PRM7	Q9prm7 xenopus. ac
13	25	22.9	19	2	Q44850	Q44850 borrelia bu
14	25	22.9	19	4	Q9UJI8	Q9uji8 homo sapien
15	24	22.0	12	12	O12088	O12088 caprine art
16	24	22.0	12	12	O12074	O12074 caprine art
17	24	22.0	12	12	O12076	O12076 caprine art
18	24	22.0	12	12	O12078	O12078 caprine art
19	24	22.0	12	12	O12080	O12080 caprine art
20	24	22.0	12	12	O12082	O12082 caprine art
21	24	22.0	12	12	O12084	O12084 caprine art
22	24	22.0	12	12	O12086	O12086 caprine art
23	24	22.0	12	12	O12090	O12090 caprine art
24	24	22.0	12	12	O12092	O12092 caprine art
25	24	22.0	12	12	O12094	O12094 caprine art
26	24	22.0	12	12	O12106	O12106 caprine art
27	24	22.0	12	12	O12108	O12108 caprine art
28	24	22.0	12	12	O12110	O12110 caprine art
29	24	22.0	12	12	O12112	O12112 caprine art
30	24	22.0	12	12	O12114	O12114 caprine art
31	24	22.0	12	12	O12116	O12116 caprine art
32	24	22.0	12	12	O12118	O12118 caprine art
33	24	22.0	15	2	Q46963	Q46963 escherichia
34	24	22.0	15	5	Q9TXC8	Q9txc8 locusta mig
35	24	22.0	16	4	Q9UCK9	Q9uck9 homo sapien
36	24	22.0	16	4	Q9UCH1	Q9uch1 homo sapien
37	24	22.0	19	2	Q47079	Q47079 escherichia
38	24	22.0	19	10	Q9S8W5	Q9s8w5 avena sativ
39	24	22.0	19	12	O93210	O93210 porcine cir
40	23	21.1	11	5	Q9V7K6	Q9v7k6 drosophila
41	23	21.1	11	7	O77892	O77892 oreochromis
42	23	21.1	11	7	O77906	O77906 oreochromis
43	23	21.1	11	7	O77918	O77918 pseudotroph
44	23	21.1	13	4	Q9Y674	Q9y674 homo sapien
45	23	21.1	17	6	Q9TR78	Q9tr78 didelphis m

ALIGNMENTS

RESULT 1
Q9PRZ1
ID Q9PRZ1 PRELIMINARY; PRT; 10 AA.
AC Q9PRZ1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE LYSYL-BRADYKININ=VASOACTIVE PEPTIDE KALLIDIN HOMOLOG.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
RN [1]
RP SEQUENCE.
RX MEDLINE; 94039817.
RA Conlon J.M., Olson K.R.;
RT "Purification of a vasoactive peptide related to lysyl-bradykinin from
RT trout plasma.";
RL FEBS Lett. 334:75-78(1993).
SQ SEQUENCE 10 AA; 1193 MW; 33C59075A3786777 CRC64;

Query Match 26.6%; Score 29; DB 13; Length 10;
Best Local Similarity 66.7%; Pred. No. 2.2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       8 EREPGW 13
           :| |||
Db       1 KRPPGW 6
```

RESULT 2
Q9UC71
ID Q9UC71 PRELIMINARY; PRT; 20 AA.
AC Q9UC71;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE FILAGGRIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95337087.
RA   Thulin C.D., Walsh K.A.;
RT   "Identification of the amino terminus of human filaggrin using
RT   differential LC/MS techniques: implications for profilaggrin
RT   processing.";
RL   Biochemistry 34:8687-8692(1995).
SQ   SEQUENCE   20 AA;  2199 MW;  0FD0856B199AB332 CRC64;


  Query Match             26.6%;  Score 29;  DB 4;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 4.6e+02;
  Matches    5;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      4 HLQGEREPGW 13
          | | |   ||
Db      5 HEQSESSHGW 14


RESULT   3
Q9PRJ4
ID   Q9PRJ4      PRELIMINARY;      PRT;     9 AA.
AC   Q9PRJ4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   BRADYKININ.
OS   Lepisosteus osseus (Long-nosed gar), and Amia calva (Bowfin).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
OC   Lepisosteus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95380361.
RA   Conlon J.M., Platzack B., Marra L.E., Youson J.H., Olson K.R.;
RT   "Isolation and biological activity of [Trp5]bradykinin from the plasma
RT   of the phylogenetically ancient fish, the bowfin and the longnosed
RT   gar.";
RL   Peptides 16:485-489(1995).
SQ   SEQUENCE   9 AA;  1099 MW;  3393D775A3786777 CRC64;


  Query Match             25.7%;  Score 28;  DB 13;  Length 9;
  Best Local Similarity   80.0%;  Pred. No. 3e+05;
  Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      9 REPGW 13
          | |||
Db      1 RPPGW 5


RESULT   4
O18502
ID   O18502      PRELIMINARY;      PRT;    14 AA.
AC   O18502;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   LYSOPHOSPHOLIPASE HOMOLOG (FRAGMENT).
GN   SMLPLH.
OS   Schistosoma mansoni (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGYPTIAN;
RA   Hamdan F.F., Ribeiro P.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF006679; AAC62255.1; -.
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1541 MW;  48B847C2E5D89177 CRC64;


  Query Match             25.7%;  Score 28;  DB 5;  Length 14;
  Best Local Similarity   38.5%;  Pred. No. 4.5e+02;
  Matches    5;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGEREPGW 13
          : ||  |:   ||
Db      2 IFLHGLGDTGHGW 14


RESULT   5
Q9R4W5
ID   Q9R4W5      PRELIMINARY;      PRT;    20 AA.
AC   Q9R4W5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   30 KDA MAJOR HEAT SHOCK PROTEIN (FRAGMENT).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 95020803.
RA   Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,
RA   Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
RT   "Heat shock protein produced by Helicobacter pylori.";
RL   Microbiol. Immunol. 38:403-405(1994).
DR   INTERPRO; IPR002026; -.
DR   PFAM; PF00547; urease_gamma; 1.
SQ   SEQUENCE   20 AA;  2302 MW;  29C9DFBFD6D21805 CRC64;


  Query Match             24.8%;  Score 27;  DB 2;  Length 20;
  Best Local Similarity   62.5%;  Pred. No. 9.4e+02;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGE 8
          |:||  ||
Db     11 LMLHYAGE 18


RESULT   6
O77893
ID   O77893      PRELIMINARY;      PRT;    11 AA.
AC   O77893;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MHC CLASS II B LOCUS 10 (FRAGMENT).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Acanthopterygii;
OC   Percomorpha; Perciformes; Labroidei; Cichlidae; Oreochromis.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 98315113.
RA   Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA   Sueltmann H., Figueroa F., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT   class II B loci.";
RL   Genetics 149:1527-1547(1998).
DR   EMBL; AF050003; AAC41342.1; -.
KW   MHC.
FT   NON_TER       1      1
FT   NON_TER      11     11
SQ   SEQUENCE   11 AA;  1296 MW;  68775B73786B572B CRC64;


  Query Match             23.9%;  Score 26;  DB 7;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 7.3e+02;
  Matches    3;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      13 WIKQLF 18
           ||::|:
Db       1 WIRELY 6


RESULT   7
Q9X3I3
ID   Q9X3I3      PRELIMINARY;      PRT;    17 AA.
AC   Q9X3I3;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   CYTOCHROME B (FRAGMENT).
GN   PETB.
OS   Prochlorococcus sp.
OC   Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC   Prochlorococcus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urbach E., Chisholm S.W.;
RT   "Genetic diversity in Prochlorococcus populations flow cytometrically
RT   sorted from the Sargasso Sea and Gulf Stream.";
RL   Limnol. Oceanog. 43:1615-1630(1998).
DR   EMBL; AF070176; AAD20791.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   17 AA;  1953 MW;  630843039ADD51B4 CRC64;


  Query Match             23.9%;  Score 26;  DB 2;  Length 17;
  Best Local Similarity   33.3%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 LLLHLQGEREPG 12
          :|:|   :|: |
Db      1 MLMHFLXDRKQG 12


RESULT   8
Q9T2R9
ID   Q9T2R9      PRELIMINARY;      PRT;    17 AA.
AC   Q9T2R9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOCHROME-C REDUCTASE 55 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS   Solanum tuberosum (Potato).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RA   Braun H.P., Kruft V., Schmitz U.K.;
RL   Planta 193:99-106(1994).
SQ   SEQUENCE   17 AA;  1758 MW;  F95F7BF7940F5F21 CRC64;


  Query Match             23.9%;  Score 26;  DB 8;  Length 17;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 LLLHLQG 7
          |:||: |
Db      4 LMLHIDG 10


RESULT   9
Q9UCG3
ID   Q9UCG3      PRELIMINARY;      PRT;    19 AA.
AC   Q9UCG3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ALPHA 2-PLASMIN INHIBITOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93249387.
RA   Bangert K., Johnsen A.H., Christensen U., Thorsen S.;
RT   "Different N-terminal forms of alpha 2-plasmin inhibitor in human
RT   plasma.";
RL   Biochem. J. 291:623-625(1993).
SQ   SEQUENCE   19 AA;  2065 MW;  41352BF04D1EEAE9 CRC64;


  Query Match             23.4%;  Score 25.5;  DB 4;  Length 19;
  Best Local Similarity   63.6%;  Pred. No. 1.5e+03;
  Matches    7;  Conservative    1;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      2 LLHLQGEREPG 12
          || | | :|||
Db     10 LLKL-GNQEPG 19


RESULT  10
Q16045
ID   Q16045      PRELIMINARY;      PRT;    14 AA.
AC   Q16045;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   D3 DOPAMINE RECEPTOR (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93326145.
RA   Nagai Y., Ueno S., Saeki Y., Soga F., Yanagihara T.;
RT   "Expression of the D3 dopamine receptor gene and a novel variant
RT   transcript generated by alternative splicing in human peripheral blood
RT   lymphocytes.";
RL   Biochem. Biophys. Res. Commun. 194:368-374(1993).
DR   EMBL; S63845; CAB32270.1; -.
FT   NON_TER       1      1
FT   NON_TER      14     14
SQ   SEQUENCE   14 AA;  1586 MW;  EA310BEFE94CF1B1 CRC64;


  Query Match             22.9%;  Score 25;  DB 4;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+03;
  Matches    5;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 LLHLQ 6
          |||||
Db      9 LLHLQ 13


RESULT  11
Q9R5F6
ID   Q9R5F6      PRELIMINARY;      PRT;    18 AA.
AC   Q9R5F6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   UREASE SMALL SUBUNIT (FRAGMENT).
OS   Helicobacter mustelae.
OC   Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC   Helicobacter.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93084378.
RA   Turbett G.R., Hoj P.B., Horne R., Mee B.J.;
RT   "Purification and characterization of the urease enzymes of
```

```
RT   Helicobacter species from humans and animals.";
RL   Infect. Immun. 60:5259-5266(1992).
SQ   SEQUENCE   18 AA;  2060 MW;  29C8E0AB77E21805 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 18;
  Best Local Similarity   50.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 LLLHLQGE 8
            ::||  ||
Db       9 MMLHYAGE 16


RESULT  12
Q9PRM7
ID   Q9PRM7      PRELIMINARY;      PRT;    18 AA.
AC   Q9PRM7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ACID THIOL PROTEASE (FRAGMENT).
OS   Xenopus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96417001.
RA   Miyata S., Kihara H.K.;
RT   "Cathepsin L-like protease from Xenopus embryos that is stimulated by
RT   nucleoside phosphates and nucleic acids.";
RL   Zool. Sci. 12:771-774(1995).
SQ   SEQUENCE   18 AA;  2060 MW;  A1DC9B106B341504 CRC64;


  Query Match             22.9%;  Score 25;  DB 13;  Length 18;
  Best Local Similarity   57.1%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       9 REPGWIK 15
            || |::|
Db       8 REKGYVK 14


RESULT  13
Q44850
ID   Q44850      PRELIMINARY;      PRT;    19 AA.
AC   Q44850;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   HYPOTHETICAL PROTEIN (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RA   Akins D.R., Popova T., Brusca J., Goldberg M.L., Li M., Baker S.C.,
RA   Norgard M.V., Radolf J.D.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; L31425; AAA64901.1; -.
KW   Hypothetical protein.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2321 MW;  C821BF312DB41512 CRC64;


  Query Match             22.9%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   50.0%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 LLHLQGEREP 11
            |:|||    |
Db       4 LIHLQSILHP 13


RESULT  14
Q9UJI8
ID   Q9UJI8      PRELIMINARY;      PRT;    19 AA.
AC   Q9UJI8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   DJ333A15.2 (PROSTAGLANDIN E RECEPTOR (EP3E)) (FRAGMENT).
GN   PTGER3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Frankland J.;
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL031429; CAB52457.1; -.
KW   Receptor.
FT   NON_TER       1      1
SQ   SEQUENCE   19 AA;  2245 MW;  142450379B298FEC CRC64;


  Query Match             22.9%;  Score 25;  DB 4;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.8e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 LHLQGEREP 11
            ||: |: :|
Db       8 LHIVGKYKP 16


RESULT  15
O12088
ID   O12088      PRELIMINARY;      PRT;    12 AA.
AC   O12088;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   TAT PROTEIN (FRAGMENT).
GN   TAT.
OS   Caprine arthritis encephalitis virus (CAEV).
OC   Viruses; Retroid viruses; Retroviridae; Lentivirus.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Turelli P., Guiguen F., Mornex J.F., Vigne R., Querat G.;
RL   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U81435; AAB60824.1; -.
FT   NON_TER       1      1
SQ   SEQUENCE   12 AA;  1279 MW;  4B90BBB1E8644EB7 CRC64;


  Query Match             22.0%;  Score 24;  DB 12;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 1.6e+03;
  Matches    3;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      11 PGW 13
           |||
Db       8 PGW 10


Search completed: December 21, 2000, 08:35:40
Job time: 603 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:31:50 ; Search time 99.91 Seconds
(without alignments)
3.355 Million cell updates/sec

Title: US-08-934-367-37
Perfect score: 104
Sequence: 1 NITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 164575 seqs, 16761186 residues

Total number of hits satisfying chosen parameters: 87906

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	29	27.9	12	1	US-08-164-618-20	Sequence 20, Appl
2	29	27.9	14	1	US-08-159-340A-17	Sequence 17, Appl
3	27	26.0	10	2	US-08-706-702-21	Sequence 21, Appl
4	27	26.0	10	3	US-08-706-706-21	Sequence 21, Appl
5	27	26.0	13	1	US-08-306-231-9	Sequence 9, Appli
6	27	26.0	13	1	US-08-355-888A-25	Sequence 25, Appl
7	27	26.0	13	2	US-08-693-697-25	Sequence 25, Appl
8	27	26.0	13	3	US-08-693-696-25	Sequence 25, Appl
9	26	25.0	15	2	US-08-484-905-28	Sequence 28, Appl
10	26	25.0	15	3	US-08-481-985B-28	Sequence 28, Appl
11	26	25.0	15	5	5304631-8	Patent No. 5304631
12	26	25.0	15	5	5304631-16	Patent No. 5304631
13	26	25.0	16	1	US-08-218-025A-12	Sequence 12, Appl
14	26	25.0	17	2	US-08-182-067-14	Sequence 14, Appl
15	26	25.0	17	2	US-08-465-313-14	Sequence 14, Appl
16	26	25.0	17	5	5304631-5	Patent No. 5304631
17	26	25.0	20	2	US-08-466-975A-12	Sequence 12, Appl
18	26	25.0	20	2	US-08-466-975A-13	Sequence 13, Appl
19	26	25.0	20	2	US-08-391-671A-12	Sequence 12, Appl
20	26	25.0	20	2	US-08-391-671A-13	Sequence 13, Appl
21	26	25.0	20	2	US-08-475-634D-2	Sequence 2, Appli
22	26	25.0	20	3	US-08-467-902A-12	Sequence 12, Appl
23	26	25.0	20	3	US-08-467-902A-13	Sequence 13, Appl
24	25	24.0	15	2	US-08-480-190-77	Sequence 77, Appl
25	25	24.0	15	2	US-08-488-379-77	Sequence 77, Appl
26	25	24.0	15	4	PCT-US93-07545-77	Sequence 77, Appl
27	25	24.0	16	2	US-08-480-190-76	Sequence 76, Appl
28	25	24.0	16	2	US-08-488-379-76	Sequence 76, Appl
29	25	24.0	16	4	PCT-US93-07545-76	Sequence 76, Appl
30	25	24.0	17	1	US-07-894-212A-3	Sequence 3, Appli
31	25	24.0	18	1	US-07-893-928A-3	Sequence 3, Appli
32	25	24.0	18	2	US-08-480-190-192	Sequence 192, App
33	25	24.0	18	2	US-08-488-379-192	Sequence 192, App
34	25	24.0	18	4	PCT-US93-07545-192	Sequence 192, App
35	25	24.0	19	2	US-08-480-190-74	Sequence 74, Appl
36	25	24.0	19	2	US-08-480-190-75	Sequence 75, Appl
37	25	24.0	19	2	US-08-488-379-74	Sequence 74, Appl
38	25	24.0	19	2	US-08-488-379-75	Sequence 75, Appl
39	25	24.0	19	2	US-08-696-944-3	Sequence 3, Appli
40	25	24.0	19	4	PCT-US93-07545-74	Sequence 74, Appl
41	25	24.0	19	4	PCT-US93-07545-75	Sequence 75, Appl
42	25	24.0	20	2	US-08-480-190-191	Sequence 191, App
43	25	24.0	20	2	US-08-749-852-55	Sequence 55, Appl
44	25	24.0	20	2	US-08-488-379-191	Sequence 191, App
45	25	24.0	20	4	PCT-US93-07545-191	Sequence 191, App

ALIGNMENTS

```
RESULT 1
US-08-164-618-20
; Sequence 20, Application US/08164618
; Patent No. 5408036
; GENERAL INFORMATION:
;   APPLICANT: Ghadiri, M. Reza
;   TITLE OF INVENTION: Isolated Metallopolypeptides:
;   TITLE OF INVENTION: Compositions and Synthetic Methods
;   NUMBER OF SEQUENCES: 20
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &
;     ADDRESSEE: Milnamow, Ltd.
;     STREET: 180 No. 5408036th Stetson, Suite 4700
;     CITY: Chicago
;     STATE: Illinois
;     COUNTRY: USA
;     ZIP: 60601
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/164,618
;     FILING DATE:
;     CLASSIFICATION: 530
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/07/769,621
;     FILING DATE:
;     APPLICATION NUMBER: 07/591,988
;     FILING DATE: October 2, 1990
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Gamson, Edward G.
;     REGISTRATION NUMBER: 29,381
;     REFERENCE/DOCKET NUMBER: SCRF 231.0
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (312) 616-5400
;     TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 20:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 12 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-164-618-20


  Query Match 27.9%; Score 29; DB 1; Length 12;
  Best Local Similarity 71.4%; Pred. No. 40;
  Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      14 VKYGLHN 20
             ||:| ||
Db       3 VKHGSHN 9


RESULT   2
US-08-159-340A-17
; Sequence 17, Application US/08159340A
; Patent No. 5565352
;  GENERAL INFORMATION:
;    APPLICANT:  Hochstrasser, Mark
;    APPLICANT:  Papa, Feroz
;    TITLE OF INVENTION:  DEUBIQUITINATING ENZYME:  COMPOSITIONS
;    TITLE OF INVENTION:  AND METHODS
;    NUMBER OF SEQUENCES:  44
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Arnold, White & Durkee
;      STREET:  P.O. Box 4433
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  USA
;      ZIP:  77210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/159,340A
;      FILING DATE:  24-NOV-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Highlander, Steven L.
;      REGISTRATION NUMBER:  37,642
;      REFERENCE/DOCKET NUMBER:  ARCD:112/HYL
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (512) 418-3000
;      TELEFAX:  (512) 474-7577
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  14 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-159-340A-17


  Query Match             27.9%;  Score 29;  DB 1;  Length 14;
  Best Local Similarity   71.4%;  Pred. No. 48;
  Matches    5;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      14 VKYGLHN 20
             | || ||
Db       8 VHYGTHN 14


RESULT   3
US-08-706-702-21
; Sequence 21, Application US/08706702
; Patent No. 5948614
;  GENERAL INFORMATION:
;    APPLICANT:  Chatterjee, Deb K.
;    TITLE OF INVENTION:  Cloned DNA Polymerases from Thermotoga
;    TITLE OF INVENTION:  maritima and Mutants Thereof
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 New York Ave., N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/706,702
;      FILING DATE:  06-SEP-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/689,807
;      FILING DATE:  14-AUG-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/537,400
;      FILING DATE:  02-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/576,759
;      FILING DATE:  21-DEC-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/537,397
;      FILING DATE:  02-OCT-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/525,057
;      FILING DATE:  08-SEP-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Esmond, Robert W.
;      REGISTRATION NUMBER:  32,893
;      REFERENCE/DOCKET NUMBER:  0942.2800006
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-2600
;      TELEFAX:  202-371-2540
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  10 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  peptide
US-08-706-702-21


  Query Match             26.0%;  Score 27;  DB 2;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 73;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMM 9
             |||:| |:
Db       3 ITGDKDML 10


RESULT   4
US-08-706-706-21
; Sequence 21, Application US/08706706
; Patent No. 6015668
;  GENERAL INFORMATION:
;    APPLICANT:  Hughes, A. John
;    APPLICANT:  Chatterjee, Deb K.
;    TITLE OF INVENTION:  Cloned DNA Polymerases from Thermotoga and
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 New York Ave., N.W., Suite 600
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/706,706
```

```
;       FILING DATE:  06-SEP-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/689,807
;       FILING DATE:  14-AUG-1996
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/537,400
;       FILING DATE:  02-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/370,190
;       FILING DATE:  01-JAN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/316,423
;       FILING DATE:  30-SEP-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/576,759
;       FILING DATE:  21-DEC-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/537,397
;       FILING DATE:  02-OCT-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/525,057
;       FILING DATE:  08-SEP-1995
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Esmond, Robert W.
;       REGISTRATION NUMBER:  32,893
;       REFERENCE/DOCKET NUMBER:  0942.2800005
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-2600
;       TELEFAX:  202-371-2540
;   INFORMATION FOR SEQ ID NO:  21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  10 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:  not relevant
;       TOPOLOGY:  not relevant
;     MOLECULE TYPE:  peptide
US-08-706-706-21


  Query Match             26.0%;  Score 27;  DB 3;  Length 10;
  Best Local Similarity   62.5%;  Pred. No. 73;
  Matches    5;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMM 9
            |||:| |:
Db       3 ITGDKDML 10


RESULT   5
US-08-306-231-9
; Sequence 9, Application US/08306231
; Patent No. 5643748
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,231
;      FILING DATE:  14-SEP-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-076
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-306-231-9


  Query Match             26.0%;  Score 27;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 98;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   6
US-08-355-888A-25
; Sequence 25, Application US/08355888A
; Patent No. 5763211
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/355,888A
;      FILING DATE:  14-DEC-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-078
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
```

```
;    MOLECULE TYPE:  peptide
US-08-355-888A-25


  Query Match             26.0%;  Score 27;  DB 1;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 98;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   7
US-08-693-697-25
; Sequence 25, Application US/08693697
; Patent No. 5869610
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,697
;      FILING DATE:  05-AUG-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  8907-0037-999
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-493-4935
;      TELEFAX:  650-493-5556
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-693-697-25


  Query Match             26.0%;  Score 27;  DB 2;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 98;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   8
US-08-693-696-25
; Sequence 25, Application US/08693696
; Patent No. 6005080
;  GENERAL INFORMATION:
;    APPLICANT:  Snodgrass, H. R.
;    APPLICANT:  Cioffi, Joseph
;    APPLICANT:  Zupancic, Thomas J.
;    APPLICANT:  Shafer, Alan W.
;    TITLE OF INVENTION:  Hu-B1.219, A NOVEL HUMAN HEMATOPOIETIN
;    TITLE OF INVENTION:  RECEPTOR
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/693,696
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/355,888
;      FILING DATE:  14-DEC-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Poissant, Brian M.
;      REGISTRATION NUMBER:  28,462
;      REFERENCE/DOCKET NUMBER:  7225-078
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  25:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  13 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-693-696-25


  Query Match             26.0%;  Score 27;  DB 3;  Length 13;
  Best Local Similarity   66.7%;  Pred. No. 98;
  Matches    4;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      13 QVKYGL 18
           |::|||
Db       3 QIRYGL 8


RESULT   9
US-08-484-905-28
; Sequence 28, Application US/08484905
; Patent No. 5976551
;  GENERAL INFORMATION:
;    APPLICANT:  Mottez, Estelle
;    APPLICANT:  Abastado, Jean-Pierre
;    APPLICANT:  Kourilsky, Philippe
;    TITLE OF INVENTION:  An Altered Major Histocompatibility
;    TITLE OF INVENTION:  Complex(MHC) Determinant and Methods for Using the
;    TITLE OF INVENTION:  Determinant
;    NUMBER OF SEQUENCES:  127
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W., Suite 700
;      CITY:  Washington
;      STATE:  D.C.
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy Disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS-/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/484,905
;       FILING DATE:  07-JUNE-1995
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/801,818
;       FILING DATE:  05-DEC-1991
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/792,473
;       FILING DATE:  15-NOV-1991
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Potter, Jane E. R.
;       REGISTRATION NUMBER:  33,332
;       REFERENCE/DOCKET NUMBER:  03495.0106-03000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-408-4000
;       TELEFAX:  202-408-4400
;   INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-484-905-28


  Query Match             25.0%;  Score 26;  DB 2;  Length 15;
  Best Local Similarity   38.5%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLGQVK 15
           | |    :|||::
Db       3 TSEGCRQILGQLQ 15


RESULT  10
US-08-481-985B-28
; Sequence 28, Application US/08481985B
; Patent No. 6011146
;  GENERAL INFORMATION:
;     APPLICANT:  Mottez, Estelle
;     APPLICANT:  Abastado, Jean-Pierre
;     APPLICANT:  Kourilsky, Phillipe
;     TITLE OF INVENTION:  Altered Major Histocompatibility Complex
;     TITLE OF INVENTION:
;     NUMBER OF SEQUENCES:  148
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;       ADDRESSEE:  Dunner
;       STREET:  1300 I Street, N.W., Suite 700
;       CITY:  Washington
;       STATE:  D.C.
;       ZIP:  20005-3315
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/481,985B
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/801,818
;       FILING DATE:  05-DEC-1991
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/792,473
;       FILING DATE:  15-NOV-1991
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Meyers, Kenneth J.
;       REGISTRATION NUMBER:  25,146
;       REFERENCE/DOCKET NUMBER:  03495.0106-04000
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-408-4000
;       TELEFAX:  202-408-4400
;   INFORMATION FOR SEQ ID NO:  28:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  15 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-481-985B-28


  Query Match             25.0%;  Score 26;  DB 3;  Length 15;
  Best Local Similarity   38.5%;  Pred. No. 1.7e+02;
  Matches    5;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLGQVK 15
           | |    :|||::
Db       3 TSEGCRQILGQLQ 15


RESULT  11
5304631-8
;Patent No. 5304631
;     APPLICANT: STEWART, JOHN M.;HAHN, KARL W.;. KLIS, WIESLAW A.
;     TITLE OF INVENTION: SYNTHETIC HELIZYME ENZYMES
;     NUMBER OF SEQUENCES: 16
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/464,932
;       FILING DATE: 16-JAN-1990
;SEQ ID NO:8:
;       LENGTH: 15
5304631-8


  Query Match             25.0%;  Score 26;  DB 5;  Length 15;
  Best Local Similarity   54.5%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLGQVK 15
           |||  || ::|
Db       2 EKAKKLLEELK 12


RESULT  12
5304631-16
;Patent No. 5304631
;     APPLICANT: STEWART, JOHN M.;HAHN, KARL W.;. KLIS, WIESLAW A.
;     TITLE OF INVENTION: SYNTHETIC HELIZYME ENZYMES
;     NUMBER OF SEQUENCES: 16
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/464,932
;       FILING DATE: 16-JAN-1990
;SEQ ID NO:16:
;       LENGTH: 15
5304631-16


  Query Match             25.0%;  Score 26;  DB 5;  Length 15;
  Best Local Similarity   54.5%;  Pred. No. 1.7e+02;
  Matches    6;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLGQVK 15
           |||  || ::|
Db       2 EKAKKLLEELK 12
```

```
RESULT  13
US-08-218-025A-12
; Sequence 12, Application US/08218025A
; Patent No. 5556744
;  GENERAL INFORMATION:
;    APPLICANT:  Weiner, David B.
;    APPLICANT:  Ugen, Kenneth E.
;    APPLICANT:  Williams, William V.
;    TITLE OF INVENTION:  Methods and Compositions for Diagnosing
;    TITLE OF INVENTION:  and Treating Certain HIV Infected Patients
;    NUMBER OF SEQUENCES:  197
;    CORRESPONDENCE ADDRESS:.
;      ADDRESSEE:  Howson and Howson
;      STREET:  P.O. Box 457, 321 No. 5556744ristown Road
;      CITY:  Spring House
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19477
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/218,025A
;      FILING DATE:  24-MAR-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/891,451
;      FILING DATE:  29-MAY-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bak, Mary E.
;      REGISTRATION NUMBER:  31,215
;      REFERENCE/DOCKET NUMBER:  WST33A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 540-9206
;      TELEFAX:  (215) 540-5818
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  16 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-218-025A-12


  Query Match             25.0%;  Score 26;  DB 1;  Length 16;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+02;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NITGEKAMM 9
            |:|||   |
Db       8 NVTGENFNM 16


RESULT  14
US-08-182-067-14
; Sequence 14, Application US/08182067
; Patent No. 5985279
;  GENERAL INFORMATION:
;    APPLICANT:  WALDMANN, HERMAN
;    APPLICANT:  SIMS, MARTIN
;    APPLICANT:  CROWE, SCOTT
;    TITLE OF INVENTION:  HUMANIZED ANTIBODY AGAINST CD18
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Rothwell, Figg Ernst & Kurz
;      STREET:  Suite 701-E, 555 Thirteenth St., N.W
;      CITY:  Washington
;      STATE:  D. C.
;      COUNTRY:  U.S.A.
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/182,067
;      FILING DATE:  23-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01289
;      FILING DATE:  15-JUL-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  GB 9115364.3
;      FILING DATE:  16-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ERNST, BARBARA G.
;      REGISTRATION NUMBER:  30,377
;      REFERENCE/DOCKET NUMBER:  1786-118A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)783-6040
;      TELEFAX:  (202)783-6031
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-182-067-14


  Query Match             25.0%;  Score 26;  DB 2;  Length 17;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GQVKYG 17
             |: |||
Db       7 GETKYG 12


RESULT  15
US-08-465-313-14
; Sequence 14, Application US/08465313
; Patent No. 5997867
;  GENERAL INFORMATION:
;    APPLICANT:  WALDMANN, HERMAN
;    APPLICANT:  SIMS, MARTIN J.
;    APPLICANT:  CROWE, J. SCOTT
;    TITLE OF INVENTION:  HUMANIZED ANTIBODY AGAINST CD18
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;      STREET:  TWO MILITIA DRIVE
;      CITY:  LEXINGTON
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  02173
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,313
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/182,067
;      FILING DATE:  23-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB92/01289
;      FILING DATE:  15-JUL-1992
;    PRIOR APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  GB 9115364.3
;      FILING DATE:  16-JUL-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BROOK, DAVID E.
;      REGISTRATION NUMBER:  22,592
;      REFERENCE/DOCKET NUMBER:  LYNX91-01A2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 861-6240
;      TELEFAX:  (617) 861-9540
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  17 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-465-313-14


  Query Match             25.0%;  Score 26;  DB 2;  Length 17;
  Best Local Similarity   66.7%;  Pred. No. 2e+02;
  Matches    4;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GQVKYG 17
           |: |||
Db       7 GETKYG 12


Search completed: December 21, 2000, 08:31:50
Job time: 373 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:30:08 ; Search time 112.59 Seconds
(without alignments)
11.273 Million cell updates/sec

Title: US-08-934-367-37
Perfect score: 104
Sequence: 1 NITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 182106 seqs, 63460219 residues

Total number of hits satisfying chosen parameters: 3930

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_65:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	26	25.0	18	2	I52623	hypothetical prote
2	24	23.1	15	2	PA0091	methionine adenosy
3	24	23.1	15	2	PC4213	bphB protein - Com
4	24	23.1	16	2	PH1475	T-cell receptor be
5	23	22.1	9	2	A11497	transaldolase (EC
6	23	22.1	10	2	S74147	glyceraldehyde-3-p
7	23	22.1	15	2	PH0797	T-cell receptor al
8	23	22.1	20	2	S04961	malate dehydrogena
9	23	22.1	20	2	S04958	malate dehydrogena
10	23	22.1	20	2	B44835	dTDPglucose 4,6-de
11	23	22.1	20	2	PH1338	Ig heavy chain DJ
12	22	21.2	12	2	S23168	Z protein - guinea
13	22	21.2	14	2	JN0390	histamine-releasin
14	22	21.2	14	2	S72217	D-arabinose 1-dehy
15	22	21.2	15	2	A49887	transcription fact
16	22	21.2	16	2	S02455	DNA-invertase - ph
17	22	21.2	16	2	PH0773	T-cell receptor be
18	22	21.2	16	2	PH1476	T-cell receptor be
19	22	21.2	18	2	S63487	fibrinogenolytic p
20	22	21.2	20	2	A47105	dystroglycan - chi
21	22	21.2	20	2	S10876	hypothetical prote
22	21	20.2	16	2	D45193	zinc finger protei
23	21	20.2	18	2	A41877	LcrKc - Yersinia p
24	21	20.2	18	2	A49857	megakaryocyte pote
25	21	20.2	19	2	B61056	cytochrome P450 PB
26	21	20.2	20	2	PH1358	Ig heavy chain DJ
27	20	19.2	9	2	G58502	kidney and bladder
28	20	19.2	12	2	PH1635	Ig H chain V-D-J r
29	20	19.2	13	2	PS0325	tetrahydroberberin
30	20	19.2	13	2	PH1772	T cell receptor al
31	20	19.2	14	2	PH1332	Ig heavy chain DJ
32	20	19.2	14	2	A28018	very late antigen-
33	20	19.2	15	2	PL0143	carbon-monoxide de
34	20	19.2	15	2	A30330	neuropeptide pep -
35	20	19.2	15	2	PH1619	Ig H chain V-D-J r
36	20	19.2	15	2	PH0775	T-cell receptor al
37	20	19.2	15	2	PH0779	T-cell receptor al
38	20	19.2	15	2	PH1455	T-cell receptor al
39	20	19.2	15	2	PH0752	T-cell receptor be
40	20	19.2	18	2	A39997	group III allergen
41	20	19.2	19	2	S43657	hsp90 protein homo
42	20	19.2	20	2	S11416	ribosomal protein
43	19.5	18.8	15	2	PA0106	protein QF200076 -
44	19.5	18.8	19	2	A39504	octamer-binding pr
45	19.5	18.8	20	2	C56385	nitrophorin 3 - Rh

ALIGNMENTS

RESULT 1
I52623
hypothetical protein TCR delta [imported] - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change 20-Jun-2000
C;Accession: I52623
R;Przybylski, G.; Oettle, H.; Ludwig, W.D.; Siegert, W.; Schmidt, C.A.
Br. J. Haematol. 87, 301-307, 1994
A;Title: Molecular characterization of illegitimate TCR delta gene rearrangements in
A;Reference number: I52623; MUID:95034253
A;Accession: I52623
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-18 <RES>
A;Cross-references: GB:S73537; NID:g688155
C;Genetics:
A;Gene: TCRdelta

Query Match 25.0%; Score 26; DB 2; Length 18;
Best Local Similarity 41.2%; Pred. No. 4.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 8; Gaps 1;

```
Qy      11 LGQVKY--------GLH 19
           ||| ||        |:|
Db       1 LGQKKYSSYLWGTSGVH 17
```

RESULT 2
PA0091
methionine adenosyltransferase (EC 2.5.1.6) 2 - fungus (Fusarium sporotrichioides) (f
C;Species: Fusarium sporotrichioides
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 05-May-2000
C;Accession: PA0091
R;Chow, L.P.; Fukaya, N.; Sugiura, Y.; Ueno, Y.; Tabuchi, K.; Tsugita, A.
submitted to JIPID, October 1994
A;Reference number: PA0051
A;Accession: PA0091
A;Molecule type: protein
A;Residues: 1-15 <CHO>
C;Keywords: S-adenosylmethionine; transferase

Query Match 23.1%; Score 24; DB 2; Length 15;
Best Local Similarity 60.0%; Pred. No. 8.1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 TGEKAMMLLG 12
           | |||:  ||
Db       1 TYEKALEQLG 10
```

```
RESULT   3
PC4213
bphB protein - Comamonas testosteroni (fragment)
C;Species: Comamonas testosteroni
C;Date: 17-Dec-1996 #sequence_revision 21-Jan-1997 #text_change 08-Oct-1999
C;Accession: PC4213
R;Sylvestre, M.; Sirois, M.; Hurtubise, Y.; Bergeron, J.; Ahmad, D.; Shareck, F.; Barria
Gene 174, 195-202, 1996
A;Title: Sequencing of Comamonas testosteroni strain B-356-biphenyl/chlorobiphenyl dioxy
A;Reference number: JC4993; MUID:97045812
A;Accession: PC4213
A;Molecule type: DNA
A;Residues: 1-15 <SYL>
A;Cross-references: GB:U47637; NID:g1245151; PIDN:AAC44530.1; PID:g1245156
A;Experimental source: strain B-356
C;Genetics:
A;Gene: bphB


  Query Match             23.1%;  Score 24;  DB 2;  Length 15;
  Best Local Similarity   45.5%;  Pred. No. 8.1e+02;
  Matches    5;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ITGEKAMMLLG 12
           :||| |::  |
Db       3 LTGEVALITGG 13


RESULT   4
PH1475
T-cell receptor beta chain (clone 223/5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 11-Apr-1995
C;Accession: PH1475
R;Casanova, J.L.; Martinon, F.; Gournier, H.; Barra, C.; Pannetier, C.; Regnault, A.; Kc
J. Exp. Med. 177, 811-820, 1993
A;Title: T cell receptor selection by and recognition of two class I major histocompatib
A;Reference number: PH1430; MUID:93171821
A;Accession: PH1475
A;Molecule type: mRNA
A;Residues: 1-16 <CAS>
A;Experimental source: cytolytic T-lymphocyte
C;Superfamily: immunoglobulin homology
C;Keywords: receptor; T-cell


  Query Match             23.1%;  Score 24;  DB 2;  Length 16;
  Best Local Similarity   40.0%;  Pred. No. 8.7e+02;
  Matches    4;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 TGEKAMMLLG 12
           |||:  :  |
Db       7 TGERERLFFG 16


RESULT   5
A11497
transaldolase (EC 2.2.1.2) III - yeast (Pichia jadinii) (fragment)
C;Species: Pichia jadinii, Candida utilis
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 30-Sep-1993
C;Accession: A11497
R;Tsolas, O.; Sun, S.C.
Arch. Biochem. Biophys. 167, 525-533, 1975
A;Title: Isolation of a peptide containing a histidinyl-cysteinyl sequence from the acti
A;Reference number: A11497; MUID:75145197
A;Accession: A11497
A;Molecule type: protein
A;Residues: 1-9 <TSO>
C;Keywords: transferase


  Query Match             22.1%;  Score 23;  DB 2;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 1.8e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLH 19
           ||:|
Db       1 YGIH 4


RESULT   6
S74147
glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S74147
R;Fukuda, A.; Osawa, T.; Hitomi, K.; Uchida, K.
Arch. Biochem. Biophys. 333, 419-426, 1996
A;Title: 4-Hydroxy-2-nonenal cytotoxicity in renal proximal tubular cells: protein mo
A;Reference number: S74147; MUID:96404942
A;Accession: S74147
A;Molecule type: protein
A;Residues: 1-10 <FUK>
A;Experimental source: LLC-PK1 cells (renal tubular epithelial cells)
C;Keywords: NAD; oxidoreductase


  Query Match             22.1%;  Score 23;  DB 2;  Length 10;
  Best Local Similarity   50.0%;  Pred. No. 7.8e+02;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      12 GQVKYGLH 19
           |:|| |::
Db       1 GKVKVGVN 8


RESULT   7
PH0797
T-cell receptor alpha chain (PF2.10.1 V-alpha-3.AR5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0797
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility comple
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846
A;Accession: PH0797
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: EMBL:X60903
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor


  Query Match             22.1%;  Score 23;  DB 2;  Length 15;
  Best Local Similarity   33.3%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 NITGEKAMMLLG 12
           :|||    :: |
Db       4 SITGNTRKLIFG 15


RESULT   8
S04961
malate dehydrogenase (EC 1.1.1.37) - Kibdelosporangium aridum (fragment)
C;Species: Kibdelosporangium aridum
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Aug-1994
C;Accession: S04961
R;Rommel, T.O.; Hund, H.K.; Speth, A.R.; Lingens, F.
Biol. Chem. Hoppe-Seyler 370, 763-768, 1989
A;Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydro
A;Reference number: S04956; MUID:89374824
A;Accession: S04961
A;Molecule type: protein
```

A;Residues: 1-20 <ROM>
C;Superfamily: L-lactate dehydrogenase
C;Keywords: oxidoreductase; tricarboxylic acid cycle

Query Match 22.1%; Score 23; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20
```

RESULT 9
S04958
malate dehydrogenase (EC 1.1.1.37) - Microtetraspora glauca (fragment)
C;Species: Microtetraspora glauca
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Aug-1994
C;Accession: S04958
R;Rommel, T.O.; Hund, H.K.; Speth, A.R.; Lingens, F.
Biol. Chem. Hoppe-Seyler 370, 763-768, 1989
A;Title: Purification and N-terminal amino-acid sequences of bacterial malate dehydrogen
A;Reference number: S04956; MUID:89374824
A;Accession: S04958
A;Molecule type: protein
A;Residues: 1-20 <ROM>
C;Superfamily: L-lactate dehydrogenase
C;Keywords: oxidoreductase; tricarboxylic acid cycle

Query Match 22.1%; Score 23; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      12 GQVKYGL 18
           ||: | |
Db      11 GQIGYAL 17
```

RESULT 10
B44835
dTDPglucose 4,6-dehydratase (EC 4.2.1.46) - Streptomyces peucetius (fragment)
C;Species: Streptomyces peucetius
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: B44835
R;Thompson, M.W.; Strohl, W.R.; Floss, H.G.
J. Gen. Microbiol. 138, 779-786, 1992
A;Title: Purification and characterization of TDP-D-glucose 4,6-dehydratase from anthrac
A;Reference number: A44835; MUID:92268857
A;Accession: B44835
A;Molecule type: protein
A;Residues: 1-20 <THO>
A;Cross-references: PIDN:AAB22196.1; PID:g249564
A;Experimental source: ATCC 29050
A;Note: sequence extracted from NCBI backbone (NCBIP:103673)
C;Keywords: carbon-oxygen lyase; hydro-lyase

Query Match 22.1%; Score 23; DB 2; Length 20;
Best Local Similarity 57.1%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      12 GQVKYGL 18
           ||: | |
Db      14 GQIGYAL 20
```

RESULT 11
PH1338
Ig heavy chain DJ region (clone C372-115) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: PH1338
R;Wasserman, R.; Galili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.
J. Exp. Med. 176, 1577-1581, 1992
A;Title: Predominance of fetal type DJH joining in young children with B precursor ly
A;Reference number: PH1302; MUID:93094761
A;Accession: PH1338
A;Molecule type: DNA
A;Residues: 1-20 <WAS>
C;Keywords: heterotetramer; immunoglobulin

Query Match 22.1%; Score 23; DB 2; Length 20;
Best Local Similarity 30.0%; Pred. No. 1.6e+03;
Matches 3; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

```
Qy       8 MMLLGQVKYG 17
           :: :| ::||
Db       4 ILRIGLLRYG 13
```

RESULT 12
S23168
Z protein - guinea pig (fragment)
C;Species: Cavia porcellus (guinea pig)
C;Date: 19-Mar-1997 #sequence_revision 24-Mar-1999 #text_change 24-Mar-1999
C;Accession: S23168
R;Nicollier, M.; Roblin, S.; Cypriani, B.; Remy-Martin, J.P.; Adessi, G.L.
Eur. J. Biochem. 205, 1137-1144, 1992
A;Title: Purification and characterization of a binding protein related to the Z clas
A;Reference number: S23168; MUID:92249319
A;Accession: S23168
A;Molecule type: protein
A;Residues: 1-12 <NIC>
A;Experimental source: liver
C;Function:
A;Description: binds dehydroepiandrosterone sulfate, estrone sulfate, oleic acid, cho
C;Keywords: liver; steroid binding

Query Match 21.2%; Score 22; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 1.4e+03;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 NITGE 5
           | |||
Db       2 NFTGE 6
```

RESULT 13
JN0390
histamine-releasing peptide II - oriental hornet
N;Alternate names: venom protein HR-2
C;Species: Vespa orientalis (oriental hornet)
C;Date: 15-Jan-1993 #sequence_revision 15-Jan-1993 #text_change 23-Aug-1997
C;Accession: JN0390; S10919
R;Miroshnikov, A.I.; Snezhkova, L.G.; Nazimov, I.V.; Reshetova, O.I.; Rozinov, B.V.;
Bioorg. Khim. 7, 1467-1477, 1981
A;Title: Structure and properties of histamine releasing peptides from the venom of V
A;Reference number: JN0389
A;Accession: JN0390
A;Molecule type: protein
A;Residues: 1-14 <MIR>
R;Tuichibaev, M.U.; Akhmedova, N.U.; Kazakov, I.; Korneev, A.S.; Gagel'gans, A.I.
Biochemistry (N.Y.) 53, 183-190, 1988
A;Title: Low-molecular-weight peptides of venom of the giant hornet Vespa orientalis.
A;Reference number: S06445
A;Accession: S10919
A;Molecule type: protein
A;Residues: 1-14 <TUI>
C;Superfamily: crabrolin
C;Keywords: amidated carboxyl end; venom
F;14/Modified site: amidated carboxyl end (Leu) #status experimental

```
  Query Match             21.2%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 MLLGQVKYGL 18
           ::||::  ||
Db       4 LILGKLVKGL 13


RESULT  14
S72217
D-arabinose 1-dehydrogenase (NAD(P)+) (EC 1.1.1.117) - yeast (Candida albicans) (fragmen
C;Species: Candida albicans
C;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 13-Mar-1998
C;Accession: S72217
R;Kim, S.T.; Huh, W.K.; Kim, J.Y.; Hwang, S.W.; Kang, S.O.
Biochim. Biophys. Acta 1297, 1-8, 1996
A;Title: D-arabinose dehydrogenase and biosynthesis of erythroascorbic acid in Candida a
A;Reference number: S72217; MUID:96439039
A;Accession: S72217
A;Molecule type: protein
A;Residues: 1-14 <KIM>
C;Keywords: oxidoreductase


  Query Match             21.2%;  Score 22;  DB 2;  Length 14;
  Best Local Similarity   30.8%;  Pred. No. 1.7e+03;
  Matches    4;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKYGLHN 20
           | |  :: : |:|
Db       1 MKLATEIDFXLNN 13


RESULT  15
A49887
transcription factor HNF-3 beta - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 07-May-1999
C;Accession: A49887
R;Sasaki, H.; Hogan, B.L.M.
Cell 76, 103-115, 1994
A;Title: HNF-3beta as a regulator of floor plate development.
A;Reference number: A49887; MUID:94116056
A;Accession: A49887
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-15 <SAS>
C;Keywords: alternative initiators; alternative splicing; transcription factor


  Query Match             21.2%;  Score 22;  DB 2;  Length 15;
  Best Local Similarity   55.6%;  Pred. No. 1.8e+03;
  Matches    5;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       7 AMMLLGQVK 15
           |  :|| ||
Db       4 ASSMLGAVK 12



Search completed: December 21, 2000, 08:30:08
Job time: 272 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:33:00 ; Search time 62.7 Seconds
(without alignments)
10.191 Million cell updates/sec

Title: US-08-934-367-37
Perfect score: 104
Sequence: 1 NITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 87993 seqs, 31947931 residues

Total number of hits satisfying chosen parameters: 1088

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	23	22.1	8	1	AL17_CARMA	P81820	carcinus ma
2	23	22.1	9	1	TAL3_PICJA	P17441	pichia jadi
3	23	22.1	20	1	MDH_KIBAR	P19978	kibdelospor
4	23	22.1	20	1	MDH_MICGL	P19979	microtetras
5	22	21.2	9	1	AL11_CARMA	P81814	carcinus ma
6	22	21.2	14	1	CRBL_VESOR	P17236	vespa orien
7	22	21.2	20	1	LECB_IRIHO	P36231	iris hollan
8	21	20.2	8	1	AL16_CARMA	P81819	carcinus ma
9	21	20.2	13	1	CRBL_VESTR	P17231	vespa tropi
10	21	20.2	14	1	TEMA_RANTE	P56917	rana tempor
11	21	20.2	14	1	TEMF_RANTE	P56921	rana tempor
12	20.5	19.7	13	1	IDHC_PIG	P20304	sus scrofa
13	20	19.2	15	1	DCMM_PSECA	P19920	pseudomonas
14	20	19.2	15	1	DCMM_PSECH	P19917	pseudomonas
15	20	19.2	16	1	ARCD_PSEPU	P41147	pseudomonas
16	20	19.2	19	1	PHLC_STAIN	P80924	staphylococ
17	20	19.2	19	1	UP25_UPEIN	P82031	uperoleia i
18	20	19.2	20	1	UCRQ_EQUAR	P81247	equisetum a
19	19	18.3	10	1	TKS1_AEDAE	P42634	aedes aegyp
20	19	18.3	10	1	TKS2_AEDAE	P42635	aedes aegyp
21	19	18.3	14	1	TEMC_RANTE	P56918	rana tempor
22	19	18.3	14	1	UC18_MAIZE	P80624	zea mays (m
23	19	18.3	15	1	RBS_PHYPA	P80657	physcomitre
24	19	18.3	17	1	BOL4_MEGPE	P07495	megabombus
25	19	18.3	18	1	LUXB_KRYAS	P18300	kryptophana
26	19	18.3	18	1	RL23_HALCU	P05975	halobacteri
27	19	18.3	19	1	MDH_COMAC	P80539	comamonas a
28	19	18.3	20	1	COG2_CHIOP	P34154	chionoecete
29	19	18.3	20	1	COG3_CHIOP	P34155	chionoecete
30	19	18.3	20	1	COGA_PARCM	P20732	paralithode
31	19	18.3	20	1	COGB_PARCM	P20733	paralithode
32	18	17.3	9	1	TAL1_PICJA	P17440	pichia jadi
33	18	17.3	12	1	PA2B_VIPBO	P31859	vipera beru
34	18	17.3	13	1	CRBL_VESMA	P17232	vespa manda
35	18	17.3	15	1	CDN5_LITCE	P82077	litoria cae
36	18	17.3	17	1	FLA2_BARBA	P35634	bartonella
37	18	17.3	17	1	NU4M_TRIRU	Q36834	trichophyto
38	18	17.3	19	1	RS19_SPICI	O31159	spiroplasma
39	18	17.3	20	1	ANCR_AGKBI	P33588	agkistrodon
40	18	17.3	20	1	COXQ_ONCMY	P80335	oncorhynchu
41	18	17.3	20	1	COXQ_THUOB	P80983	thunnus obe
42	17	16.3	8	1	AL15_CARMA	P81818	carcinus ma
43	17	16.3	9	1	HUTU_KLEAE	P12381	klebsiella
44	17	16.3	10	1	GON1_ALLMI	P37041	alligator m
45	17	16.3	10	1	GSI5_BACSU	P80861	bacillus su

ALIGNMENTS

```
RESULT   1
AL17_CARMA
ID   AL17_CARMA     STANDARD;      PRT;     8 AA.
AC   P81820;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 17.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       8       8       AMIDATION (POTENTIAL).
SQ   SEQUENCE   8 AA;  858 MW;  C82879D5AB46D865 CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
           ||  :||
Db       2 GQYSFGL 8


RESULT   2
TAL3_PICJA
ID   TAL3_PICJA     STANDARD;      PRT;     9 AA.
AC   P17441;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-FEB-1994 (Rel. 28, Last annotation update)
DE   TRANSALDOLASE III (EC 2.2.1.2) (FRAGMENT).
OS   Pichia jadinii (Yeast) (Candida utilis).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Pichia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 75145197.
RA   Tsolas O., Sun S.C.;
RT   "Isolation of a peptide containing a histidinyl-cysteinyl sequence
RT   from the active center of transaldolase.";
RL   Arch. Biochem. Biophys. 167:525-533(1975).
CC   -!- FUNCTION: TRANSALDOLASE IS IMPORTANT FOR THE BALANCE OF
```

```
CC       METABOLITES IN THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
CC       3-PHOSPHATE = D-ERYTHROSE 4-PHOSPHATE + D-FRUCTOSE 6-PHOSPHATE.
CC   -!- PATHWAY: NONOXYDATIVE PART OF THE PENTOSE-PHOSPHATE PATHWAY.
CC   -!- SIMILARITY: BELONGS TO THE TRANSALDOLASE FAMILY.
DR   PIR; A11497; A11497.
DR   INTERPRO; IPR001585; -.
DR   PROSITE; PS00958; TRANSALDOLASE_2; PARTIAL.
DR   PROSITE; PS01054; TRANSALDOLASE_1; PARTIAL.
KW   Transferase; Pentose shunt.
FT   NON_TER       1      1
FT   NON_TER       9      9
SQ   SEQUENCE   9 AA;  1033 MW;  325A31A44EB1E058 CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 9;
  Best Local Similarity   75.0%;  Pred. No. 8.8e+04;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     16 YGLH 19
          ||:|
Db      1 YGIH 4


RESULT   3
MDH_KIBAR
ID   MDH_KIBAR      STANDARD;      PRT;    20 AA.
AC   P19978;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Kibdelosporangium aridum.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Pseudonocardineae; Pseudonocardiaceae;
OC   Kibdelosporangium.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89374824.
RA   Rommel T.O., Hund H.-K., Speth A.R., Lingens F.;
RT   "Purification and N-terminal amino-acid sequences of bacterial malate
RT   dehydrogenases from six actinomycetales strains and from
RT   Phenylobacterium immobile, strain E.";
RL   Biol. Chem. Hoppe-Seyler 370:763-768(1989).
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   PIR; S04961; S04961.
DR   HSSP; P10584; 1BDM.
DR   INTERPRO; IPR001252; -.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  1989 MW;  79587E6F58C00AC9 CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GQVKYGL 18
          ||: | |
Db     14 GQIGYAL 20


RESULT   4
MDH_MICGL
ID   MDH_MICGL      STANDARD;      PRT;    20 AA.
AC   P19979;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MALATE DEHYDROGENASE (EC 1.1.1.37) (FRAGMENT).
GN   MDH.
OS   Microtetraspora glauca.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptosporangineae; Streptosporangiaceae;
OC   Microtetraspora.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 89374824.
RA   Rommel T.O., Hund H.-K., Speth A.R., Lingens F.;
RT   "Purification and N-terminal amino-acid sequences of bacterial malate
RT   dehydrogenases from six actinomycetales strains and from
RT   Phenylobacterium immobile, strain E.";
RL   Biol. Chem. Hoppe-Seyler 370:763-768(1989).
CC   -!- CATALYTIC ACTIVITY: L-MALATE + NAD(+) = OXALOACETATE + NADH.
CC   -!- SIMILARITY: BELONGS TO THE LDH FAMILY. MDH SUBFAMILY.
DR   PIR; S04958; S04958.
DR   HSSP; P10584; 1BDM.
DR   INTERPRO; IPR001236; -.
DR   INTERPRO; IPR001252; -.
DR   PFAM; PF00056; ldh; 1.
DR   PROSITE; PS00068; MDH; PARTIAL.
KW   Oxidoreductase; Tricarboxylic acid cycle; NAD.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2065 MW;  740129BE59D01EBE CRC64;


  Query Match             22.1%;  Score 23;  DB 1;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 8.1e+02;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GQVKYGL 18
          ||: | |
Db     11 GQIGYAL 17


RESULT   5
AL11_CARMA
ID   AL11_CARMA     STANDARD;      PRT;     9 AA.
AC   P81814;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 11.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       9      9       AMIDATION.
SQ   SEQUENCE   9 AA;  927 MW;  832D79CDCB46D861 CRC64;


  Query Match             21.2%;  Score 22;  DB 1;  Length 9;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GQVKYGL 18
          ||  :||
Db      3 GQYAFGL 9


RESULT   6
```

```
CRBL_VESOR
ID   CRBL_VESOR     STANDARD;      PRT;    14 AA.
AC   P17236;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   01-AUG-1990 (Rel. 15, Last annotation update)
DE   HISTAMINE RELEASING PEPTIDE II (HR-II).
OS   Vespa orientalis (Oriental hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Miroshnikov A.I., Snezhkova L.G., Nazimov I.V., Reshetova O.I.,
RA   Rozynov B.V., Gushchin I.S.;
RT   "Structure and properties of histamine releasing peptides from the
RT   venom of Vespa orientalis hornet.";
RL   Bioorg. Khim. 7:1467-1477(1981).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
DR   PIR; JN0390; JN0390.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1524 MW;  22015B4A6CEDFD38 CRC64;


  Query Match             21.2%;  Score 22;  DB 1;  Length 14;
  Best Local Similarity   40.0%;  Pred. No. 8.4e+02;
  Matches    4;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       9 MLLGQVKYGL 18
           ::||::  ||
Db       4 LILGKLVKGL 13


RESULT   7
LECB_IRIHO
ID   LECB_IRIHO     STANDARD;      PRT;    20 AA.
AC   P36231;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   N-ACETYL-D-GALACTOSAMINE-BINDING LECTIN SUBUNIT B (A-DISACCHARIDE-
DE   BINDING LECTIN SUBUNIT B) (FRAGMENT).
OS   Iris hollandica (Dutch iris).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Asparagales; Iridaceae; Iris.
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CV. GOLDEN HARVEST, AND CV. PROF. BLAAUW; TISSUE=BULB;
RX   MEDLINE; 94171801.
RA   Mo H., van Damme E.J.M., Peumans W.J., Goldstein I.J.;
RT   "Isolation and characterization of an
RT   N-acetyl-D-galactosamine-binding lectin from Dutch Iris bulbs which
RT   recognizes the blood group A disaccharide (GalNAc alpha 1-3Gal).";
RL   J. Biol. Chem. 269:7666-7673(1994).
CC   -!- FUNCTION: GAL / GALNAC-SPECIFIC LECTIN. AGGLUTINATES BOTH NATIVE
CC       AND TRYPSIN-TREATED RABBIT ERYTHROCYTES BUT NOT HUMAN ERYTHROCYTES
CC       IRRESPECTIVE OF BLOOD GROUP TYPE.
CC   -!- SUBUNIT: DISULFIDE-LINKED HETERODIMER OF A AND B CHAINS.
KW   Lectin.
FT   NON_TER       20      20
SQ   SEQUENCE   20 AA;  2246 MW;  DE08DA025FD17D56 CRC64;


  Query Match             21.2%;  Score 22;  DB 1;  Length 20;
  Best Local Similarity   50.0%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 EKAMMLLG 12
           |: || :|
Db       9 EETMMRIG 16


RESULT   8
AL16_CARMA
ID   AL16_CARMA     STANDARD;      PRT;     8 AA.
AC   P81819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 16.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=CEREBRAL GANGLION, AND THORACIC GANGLION;
RX   MEDLINE; 98121193.
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       8       8       AMIDATION.
SQ   SEQUENCE   8 AA;  813 MW;  7C286B45AB476878 CRC64;


  Query Match             20.2%;  Score 21;  DB 1;  Length 8;
  Best Local Similarity   57.1%;  Pred. No. 8.8e+04;
  Matches    4;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      12 GQVKYGL 18
             |   |||
Db       2 GPYSYGL 8


RESULT   9
CRBL_VESTR
ID   CRBL_VESTR     STANDARD;      PRT;    13 AA.
AC   P17231;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   VESPID CHEMOTACTIC PEPTIDE T (VESCP-T).
OS   Vespa tropica (Hornet).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC   Vespoidea; Vespidae; Vespinae; Vespa.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=VENOM;
RA   Yasuhara T., Nakajima T., Erspaer V.;
RL   (In) Sakakibara S. (eds.);
RL   Peptide chemistry 1982, pp.213-218, Protein Research Foundation,
RL   Osaka (1983).
CC   -!- FUNCTION: MAST CELL DEGRANULATING PEPTIDE. INDUCES THE CHEMOTAXIS
CC       OF NEUTROPHILS.
KW   Mast cell degranulation; Chemotaxis; Venom; Amidation.
FT   MOD_RES       13      13       AMIDATION.
SQ   SEQUENCE   13 AA;  1354 MW;  220140365DFE5338 CRC64;


  Query Match             20.2%;  Score 21;  DB 1;  Length 13;
  Best Local Similarity   44.4%;  Pred. No. 1.1e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
             :||::  ||
Db       4 ILGKILGGL 12
```

```
RESULT  10
TEMA_RANTE
ID   TEMA_RANTE     STANDARD;      PRT;    14 AA.
AC   P56917;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN A.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-POSITIVE
CC       BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1469 MW;  601653612B9DECD4 CRC64;

  Query Match             20.2%;  Score 21;  DB 1;  Length 14;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           |:|:|  |:
Db       4 LIGRVLSGI 12


RESULT  11
TEMF_RANTE
ID   TEMF_RANTE     STANDARD;      PRT;    14 AA.
AC   P56921;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TEMPORIN F.
OS   Rana temporaria (European common frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN;
RX   MEDLINE; 97175050.
RA   Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RA   Barra D.;
RT   "Temporins, antimicrobial peptides from the European red frog Rana
RT   temporaria.";
RL   Eur. J. Biochem. 242:788-792(1996).
CC   -!- FUNCTION: HAS ANTIBACTERIAL ACTIVITY AGAINST GRAM-NEGATIVE AND
CC       GRAM-POSITIVE BACTERIA.
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE BREVININ/ESCULENTIN/GAEGURIN/RUGOSIN
CC       FAMILY.
KW   Amphibian skin; Antibiotic; Amidation; Multigene family.
FT   MOD_RES       14      14       AMIDATION.
SQ   SEQUENCE   14 AA;  1441 MW;  4D1653612B9DECC3 CRC64;

  Query Match             20.2%;  Score 21;  DB 1;  Length 14;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      10 LLGQVKYGL 18
           |:|:|  |:
Db       4 LIGKVLSGI 12


RESULT  12
IDHC_PIG
ID   IDHC_PIG       STANDARD;      PRT;    13 AA.
AC   P20304;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   ISOCITRATE DEHYDROGENASE [NADP] CYTOPLASMIC (EC 1.1.1.42)
DE   (OXALOSUCCINATE DECARBOXYLASE) (IDH) (NADP+-SPECIFIC ICDH) (IDP)
DE   (FRAGMENT).
GN   IDH1.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 87308293.
RA   Bailey J.M., Colman R.F.;
RT   "Isolation of the glutamyl peptide labeled by the nucleotide analogue
RT   2-(4-bromo-2,3-dioxobutylthio)-1,N(6)-ethenoadenosine
RT   2',5'-biphosphate in the active site of NADP+-specific isocitrate
RT   dehydrogenase.";
RL   J. Biol. Chem. 262:12620-12626(1987).
RN   [2]
RP   SEQUENCE.
RX   MEDLINE; 87308292.
RA   Ehrlich R.S., Colman R.F.;
RT   "Characterization of an active site peptide modified by the substrate
RT   analogue 3-bromo-2-ketoglutarate on a single chain of dimeric NADP+-
RT   dependent isocitrate dehydrogenase.";
RL   J. Biol. Chem. 262:12614-12619(1987).
CC   -!- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) = 2-OXOGLUTARATE +
CC       CO(2) + NADPH.
CC   -!- SUBUNIT: HOMODIMER.
DR   PIR; A27371; A27371.
DR   PIR; A27372; A27372.
DR   INTERPRO; IPR001804; -.
DR   PROSITE; PS00470; IDH_IMDH; PARTIAL.
KW   Oxidoreductase; NADP; Tricarboxylic acid cycle.
FT   NON_TER       1       1
FT   ACT_SITE      5       5       POTENTIAL.
FT   NON_TER      13      13
SQ   SEQUENCE   13 AA;  1353 MW;  1B640F0E9F7C71E0 CRC64;

  Query Match             19.7%;  Score 20.5;  DB 1;  Length 13;
  Best Local Similarity   45.5%;  Pred. No. 1.4e+03;
  Matches    5;  Conservative    3;  Mismatches    2;  Indels    1;  Gaps    1;

Qy      10 LLGQVKYGLHN 20
           | |:: :|| |
Db       2 LAGEI-HGLSN 11


RESULT  13
DCMM_PSECA
ID   DCMM_PSECA     STANDARD;      PRT;    15 AA.
AC   P19920;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   CARBON MONOXIDE OXYGENASE [CYTOCHROME B-561] MEDIUM CHAIN (EC 1.2.2.4)
DE   (FRAGMENT).
OS   Pseudomonas carboxydovorans.
OC   Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC   Bradyrhizobium group; Oligotropha.
RN   [1]
```

```
RP   SEQUENCE.
RC   STRAIN=OM5;
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + FERROCYTOCHROME B-561 = CO(2) +
CC       2 H(+) + FERRICYTOCHROME B-561.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0141; PL0141.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1779 MW;  82DD3BF93E739D63 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKY 16
           ||: |   |
Db       1 MMIPGSFDY 9


RESULT  14
DCMM_PSECH
ID   DCMM_PSECH     STANDARD;      PRT;    15 AA.
AC   P19917;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-DEC-1992 (Rel. 24, Last annotation update)
DE   CARBON MONOXIDE DEHYDROGENASE MEDIUM CHAIN (EC 1.2.99.2) (FRAGMENT).
OS   Pseudomonas carboxydohydrogena.
OC   Bacteria; Proteobacteria.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 90055678.
RA   Kraut M., Hugendieck I., Herwig S., Meyer O.;
RT   "Homology and distribution of CO dehydrogenase structural genes in
RT   carboxydotrophic bacteria.";
RL   Arch. Microbiol. 152:335-341(1989).
CC   -!- CATALYTIC ACTIVITY: CO + H(2)O + ACCEPTOR = CO(2) + REDUCED
CC       ACCEPTOR.
CC   -!- COFACTOR: MOLYBDENUM.
CC   -!- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND
CC       SMALL.
DR   PIR; PL0143; PL0143.
KW   Oxidoreductase; Molybdenum.
FT   NON_TER      15     15
SQ   SEQUENCE   15 AA;  1815 MW;  90508CF93E739D63 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 15;
  Best Local Similarity   44.4%;  Pred. No. 2e+03;
  Matches    4;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       8 MMLLGQVKY 16
           ||: |   |
Db       1 MMIPGHFDY 9


RESULT  15
ARCD_PSEPU
ID   ARCD_PSEPU     STANDARD;      PRT;    16 AA.
AC   P41147;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   ARGININE/ORNITHINE ANTIPORTER (FRAGMENT).
GN   ARCD.
OS   Pseudomonas putida.
OC   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC   Pseudomonas.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 4359;
RA   Wilson S.D., Wang M., Filpula D.;
RL   Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: CATALYZES AN ELECTRONEUTRAL EXCHANGE BETWEEN ARGININE
CC       AND ORNITHINE TO ALLOW HIGH-EFFICIENCY ENERGY CONVERSION IN THE
CC       ARGININE DEIMINASE PATHWAY.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC       (POTENTIAL).
CC   -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC       PERMEASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U07185; AAA16963.1; -.
KW   Transport; Amino-acid transport; Transmembrane; Inner membrane.
FT   NON_TER      1      1
SQ   SEQUENCE   16 AA;  1644 MW;  90B48A7C8FAA9705 CRC64;


  Query Match             19.2%;  Score 20;  DB 1;  Length 16;
  Best Local Similarity   60.0%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      16 YGLHN 20
             |||::
Db       7 YGLYD 11


Search completed: December 21, 2000, 08:33:00
Job time: 443 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 21, 2000, 08:35:40 ; Search time 157.2 Seconds
(without alignments)
11.880 Million cell updates/sec

Title: US-08-934-367-37
Perfect score: 104
Sequence: 1 NITGEKAMMLLGQVKYGLHN 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 297973 seqs, 93374136 residues

Total number of hits satisfying chosen parameters: 4186

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_14:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	32	30.8	19	10	Q9S8G6	Q9s8g6 colocasia e
2	27	26.0	17	8	O03888	O03888 chlorogoniu
3	27	26.0	17	10	Q9S8U7	Q9s8u7 solanum tub
4	25	24.0	19	2	O31045	O31045 streptomyce
5	25	24.0	19	5	Q9TWD0	Q9twd0 sphaerechin
6	24.5	23.6	19	2	Q53545	Q53545 shigella so
7	24	23.1	11	5	Q9UAR8	Q9uar8 aedes aegyp
8	23	22.1	17	6	Q9TQZ5	Q9tqz5 macaca fasc
9	23	22.1	19	2	Q9R4A3	Q9r4a3 escherichia
10	23	22.1	20	2	Q9R5L1	Q9r5l1 streptomyce
11	22	21.2	8	3	Q05403	Q05403 saccharomyc
12	22	21.2	17	8	Q9T2S1	Q9t2s1 solanum tub
13	22	21.2	18	13	Q9PRM8	Q9prm8 cerastes ce
14	22	21.2	19	8	Q36277	Q36277 zea mays (m
15	22	21.2	19	10	Q9S901	Q9s901 vigna sinen
16	22	21.2	20	4	Q9UCP6	Q9ucp6 homo sapien
17	22	21.2	20	6	Q28734	Q28734 oryctolagus
18	22	21.2	20	13	Q9PRM4	Q9prm4 gallus gall
19	21	20.2	15	6	Q9TRW0	Q9trw0 oryctolagus
20	21	20.2	17	4	Q9UCF0	Q9ucf0 homo sapien
21	21	20.2	17	11	Q9QY07	Q9qy07 mus musculu
22	21	20.2	18	4	Q9UCB2	Q9ucb2 homo sapien
23	21	20.2	19	2	Q57012	Q57012 staphylococ
24	21	20.2	19	11	Q64132	Q64132 rattus norv
25	21	20.2	19	11	Q9QXP9	Q9qxp9 mus musculu
26	21	20.2	20	2	O34197	O34197 rickettsia
27	21	20.2	20	2	Q9R4I9	Q9r4i9 vibrio chol
28	21	20.2	20	10	Q9S739	Q9s739 arabidopsis
29	20	19.2	10	5	Q25355	Q25355 locusta mig
30	20	19.2	11	12	Q84247	Q84247 polyomaviru
31	20	19.2	15	5	Q9TXC8	Q9txc8 locusta mig
32	20	19.2	15	12	Q84332	Q84332 simian viru
33	20	19.2	16	4	Q9UCI8	Q9uci8 homo sapien
34	20	19.2	16	12	Q84353	Q84353 simian viru
35	20	19.2	17	2	O52748	O52748 anabaena sp
36	20	19.2	17	5	Q9TWF6	Q9twf6 artemia (br
37	20	19.2	18	11	Q9QVB0	Q9qvb0 rattus sp.
38	20	19.2	19	2	Q06140	Q06140 neisseria m
39	20	19.2	19	2	Q9R4X3	Q9r4x3 bacillus ce
40	20	19.2	19	5	Q9TWK7	Q9twk7 trypanosoma
41	20	19.2	19	12	Q84863	Q84863 unidentifie
42	20	19.2	20	2	Q9R9A5	Q9r9a5 nitrosospir
43	20	19.2	20	2	Q9R987	Q9r987 nitrosospir
44	20	19.2	20	2	Q9R5T8	Q9r5t8 campylobact
45	20	19.2	20	3	O13594	O13594 saccharomyc

ALIGNMENTS

RESULT 1
Q9S8G6
ID Q9S8G6 PRELIMINARY; PRT; 19 AA.
AC Q9S8G6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE TYPE 2 LECTIN (FRAGMENT).
OS Colocasia esculenta (Elephant's ear) (Taro).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; Liliopsida; Araceae; Colocasia.
RN [1]
RP SEQUENCE.
RX MEDLINE; 95288362.
RA Van Damme E.J., Goossens K., Smeets K., van Leuven F., Verhaert P.,
RA Peumans W.J.;
RT "The major tuber storage protein of araceae species is a lectin.
RT Characterization and molecular cloning of the lectin from Arum
RT maculatum L.";
RL Plant Physiol. 107:1147-1158(1995).
SQ SEQUENCE 19 AA; 2066 MW; F9C18865CA58608A CRC64;

Query Match 30.8%; Score 32; DB 10; Length 19;
Best Local Similarity 41.2%; Pred. No. 79;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

```
Qy       1 NITGEKAMMLLGQVKYG 17
           ||     ::  ||| ||
Db       1 NIPFTNNLLFSGQVXYG 17
```

RESULT 2
O03888
ID O03888 PRELIMINARY; PRT; 17 AA.
AC O03888;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE NADH DEHYDROGENASE (EC 1.6.99.3) (CYTOCHROME C REDUCTASE)
DE (TYPE I DEHYDROGENASE) (FRAGMENT).
GN NAD5.

```
OS   Chlorogonium elongatum.
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Haematococcaceae; Chlorogonium.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=2E;
RX   MEDLINE; 97306270.
RA   Kroymann J., Zetsche K.;
RT   "The apocytochrome-b gene in Chlorogonium elongatum
RT   (Chlamydomonadaceae): an intronic GIY-YIG ORF in green algal
RT   mitochondria.";
RL   Curr. Genet. 31:414-418(1997).
CC   -!- CATALYTIC ACTIVITY: NADH + ACCEPTOR = NAD(+) + REDUCED ACCEPTOR.
CC   -!- COFACTOR: FLAVOPROTEIN; IRON-SULFUR.
DR   EMBL; Y07814; CAA69147.1; -.
KW   Oxidoreductase; Mitochondrion.
FT   NON_TER       1       1
SQ   SEQUENCE   17 AA;  1844 MW;  BC88B2E8E529A229 CRC64;


  Query Match             26.0%;  Score 27;  DB 8;  Length 17;
  Best Local Similarity   31.2%;  Pred. No. 4.8e+02;
  Matches    5;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      3 TGEKAMMLLGQVKYGL 18
          ||:| : ::   |  |
Db      2 TGDKGLFIVHDQKVSL 17


RESULT   3
Q9S8U7
ID   Q9S8U7      PRELIMINARY;      PRT;    17 AA.
AC   Q9S8U7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   SOLANIDINE UDP-GLUCOSE GLUCOSYLTRANSFERASE, SGT=19 KDA CNBR CLEAVED
DE   FRAGMENT.
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 93044546.
RA   Stapleton A., Allen P.V., Tao H.P., Belknap W.R., Friedman M.;
RT   "Partial amino acid sequence of potato solanidine UDP-glucose
RT   glucosyltransferase purified by new anion-exchange and size exclusion
RT   media.";
RL   Protein Expr. Purif. 3:85-92(1992).
SQ   SEQUENCE   17 AA;  1783 MW;  75BADCB0077E593C CRC64;


  Query Match             26.0%;  Score 27;  DB 10;  Length 17;
  Best Local Similarity   83.3%;  Pred. No. 4.8e+02;
  Matches    5;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NITGEK 6
          ||:|||
Db      8 NISGEK 13


RESULT   4
O31045
ID   O31045      PRELIMINARY;      PRT;    19 AA.
AC   O31045;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   LEUA LEADER PEPTIDE.
GN   LEUA.
OS   Streptomyces coelicolor.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=D132;
RA   Potter C.A., Baumberg S.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF026444; AAB82585.1; -.
SQ   SEQUENCE   19 AA;  2080 MW;  1A591DC2999760D4 CRC64;


  Query Match             24.0%;  Score 25;  DB 2;  Length 19;
  Best Local Similarity   44.4%;  Pred. No. 1.2e+03;
  Matches    4;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     10 LLGQVKYGL 18
          :|| :::||
Db      1 MLGIMRFGL 9


RESULT   5
Q9TWD0
ID   Q9TWD0      PRELIMINARY;      PRT;    19 AA.
AC   Q9TWD0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   SIALOPROTEIN I (FRAGMENT).
OS   Sphaerechinus granularis (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Echinozoa; Echinoidea;
OC   Euechinoidea; Echinacea; Temnopleuroida; Toxopneustidae;
OC   Sphaerechinus.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96428846.
RA   Karamanos N.K., Manouras A., Anagnostides S., Makatsori E.,
RA   Tsegenidis T., Antonopoulos C.A.;
RT   "Isolation, biochemical and immunological characterisation of two sea
RT   urchin glycoproteins bearing sulphated poly(sialic acid)
RT   polysaccharides rich in N-glycolyl neuraminic acid.";
RL   Biochimie 78:171-182(1996).
SQ   SEQUENCE   19 AA;  2265 MW;  38FFDEB1A033E850 CRC64;


  Query Match             24.0%;  Score 25;  DB 5;  Length 19;
  Best Local Similarity   27.3%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      6 KAMMLLGQVKY 16
          : |:  |:::|
Db      8 QXMLTXGEIEY 18


RESULT   6
Q53545
ID   Q53545      PRELIMINARY;      PRT;    19 AA.
AC   Q53545;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   VIRF (FRAGMENT).
GN   VIRF.
OS   Shigella sonnei.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Shigella.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95394812.
RA   Nakayama S., Watanabe H.;
RT   "Involvement of cpxA, a sensor of a two-component regulatory system,
RT   in the pH-dependent regulation of expression of Shigella sonnei virF
RT   gene.";
RL   J. Bacteriol. 177:5062-5069(1995).
```

```
DR   EMBL; S79443; AAB35192.1; -.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2344 MW;  52A53B9DAAEEFCF4 CRC64;


  Query Match             23.6%;  Score 24.5;  DB 2;  Length 19;
  Best Local Similarity   38.9%;  Pred. No. 1.4e+03;
  Matches    7;  Conservative    2;  Mismatches    4;  Indels    5;  Gaps    1;

Qy      8 MMLLG-----QVKYGLHN 20
          || :|      :|  |||
Db      1 MMDMGHKNKIDIKVRLHN 18


RESULT   7
Q9UAR8
ID   Q9UAR8      PRELIMINARY;      PRT;    11 AA.
AC   Q9UAR8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   SIALOKININ I PREPROPROTEIN (FRAGMENT).
OS   Aedes aegypti (Yellowfever mosquito).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
OC   Culicidae; Aedes.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ROCKEFELLER/RED; TISSUE=SALIVARY GLAND;
RA   Beerntsen B.T., Champagne D.E., Coleman J.L., Campos Y.A., James A.A.;
RT   "Characterization of the sialokinin I gene encoding the salivary
RT   vasodilator of the yellow fever mosquito, Aedes aegypti.";
RL   Insect Mol. Biol. 0:0-0(1999).
DR   EMBL; AF108100; AAD16884.1; -.
DR   INTERPRO; IPR002040; -.
DR   PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
FT   NON_TER      1      1
SQ   SEQUENCE   11 AA;  1203 MW;  8BADC77C6B59C33A CRC64;


  Query Match             23.1%;  Score 24;  DB 5;  Length 11;
  Best Local Similarity   50.0%;  Pred. No. 9.4e+02;
  Matches    5;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 TGEKAMMLLG 12
          ||:|   |:|
Db      2 TGDKFYGLMG 11


RESULT   8
Q9TQZ5
ID   Q9TQZ5      PRELIMINARY;      PRT;    17 AA.
AC   Q9TQZ5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE (FRAGMENT).
OS   Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96273610.
RA   Nicolas M.G., Fujiki K., Murayama K., Suzuki M.T., Mineki R.,
RA   Hayakawa M., Yoshikawa Y., Cho F., Kanai A.;
RT   "Studies on the mechanism of early onset macular degeneration in
RT   cynomolgus (Macaca fascicularis) monkeys. I. Abnormal concentrations
RT   of two proteins in the retina.";
RL   Exp. Eye Res. 62:211-219(1996).
SQ   SEQUENCE   17 AA;  1671 MW;  52CCD0D1A98B3DAF CRC64;


  Query Match             22.1%;  Score 23;  DB 6;  Length 17;
  Best Local Similarity   50.0%;  Pred. No. 2.3e+03;
  Matches    4;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     12 GQVKYGLH 19
          |:|| |::
Db      1 GKVKVGVN 8


RESULT   9
Q9R4A3
ID   Q9R4A3      PRELIMINARY;      PRT;    19 AA.
AC   Q9R4A3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   N-ACETYL-D-NEURAMINIC ACID LYASE (EC 4.1.3.3) (FRAGMENT).
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96276413.
RA   Ferrero M.A., Reglero A., Fernandez-Lopez M., Ordas R.,
RA   Rodriguez-Aparicio L.B.;
RT   "N-acetyl-D-neuraminic acid lyase generates the sialic acid for
RT   colominic acid biosynthesis in Escherichia coli K1.";
RL   Biochem. J. 317:157-165(1996).
SQ   SEQUENCE   19 AA;  2061 MW;  72022247A20EBBE3 CRC64;


  Query Match             22.1%;  Score 23;  DB 2;  Length 19;
  Best Local Similarity   45.5%;  Pred. No. 2.6e+03;
  Matches    5;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 NITGEKAMMLL 11
          |: |   | :||
Db      3 NLRGVMAALLL 13


RESULT  10
Q9R5L1
ID   Q9R5L1      PRELIMINARY;      PRT;    20 AA.
AC   Q9R5L1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   THYMIDINE 5'-DIPHOSPHATE D-GLUCOSE 4,6-DEHYDRATASE (EC 4.2.1.46)
DE   (FRAGMENT).
OS   Streptomyces peucetius.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 92268857.
RA   Thompson M.W., Strohl W.R., Floss H.G.;
RT   "Purification and characterization of TDP-D-glucose 4,6-dehydratase
RT   from anthracycline-producing streptomycetes.";
RL   J. Gen. Microbiol. 138:779-786(1992).
SQ   SEQUENCE   20 AA;  1943 MW;  7958730238C00AC9 CRC64;


  Query Match             22.1%;  Score 23;  DB 2;  Length 20;
  Best Local Similarity   57.1%;  Pred. No. 2.7e+03;
  Matches    4;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     12 GQVKYGL 18
          ||: | |
Db     14 GQIGYAL 20


RESULT  11
Q05403
```

```
ID   Q05403      PRELIMINARY;      PRT;     8 AA.
AC   Q05403;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   DNA FOR ORF'S FROM CHROMOSOME XV (FRAGMENT).
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC   Saccharomycetaceae; Saccharomyces.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FY1679;
RX   MEDLINE; 96021609.
RA   Zumstein E., Pearson B.M., Kalogeropoulos A., Schweizer M.;
RT   "A 29.425 kb segment on the left arm of yeast chromosome XV contains
RT   more than twice as many unknown as known open reading frames.";
RL   Yeast 11:975-986(1995).
DR   EMBL; X83121; CAA58183.1; -.
FT   NON_TER       8      8
SQ   SEQUENCE   8 AA;  879 MW;  7B5322D2C441E058 CRC64;


  Query Match             21.2%;  Score 22;  DB 3;  Length 8;
  Best Local Similarity   75.0%;  Pred. No. 3e+05;
  Matches    3;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     17 GLHN 20
          |:||
Db      1 GIHN 4


RESULT  12
Q9T2S1
ID   Q9T2S1      PRELIMINARY;      PRT;    17 AA.
AC   Q9T2S1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CYTOCHROME-C REDUCTASE 55 KDA SUBUNIT (EC 1.10.2.2) (FRAGMENT).
OS   Solanum tuberosum (Potato).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Asteridae; euasterids I; Solanales;
OC   Solanaceae; Solanum.
RN   [1]
RP   SEQUENCE.
RA   Braun H.P., Kruft V., Schmitz U.K.;
RL   Planta 193:99-106(1994).
SQ   SEQUENCE   17 AA;  1870 MW;  207804E213CD4009 CRC64;


  Query Match             21.2%;  Score 22;  DB 8;  Length 17;
  Best Local Similarity   44.4%;  Pred. No. 3.3e+03;
  Matches    4;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      6 KAMMLLGQV 14
          ||: :|| :
Db      1 KAVDILGDI 9


RESULT  13
Q9PRM8
ID   Q9PRM8      PRELIMINARY;      PRT;    18 AA.
AC   Q9PRM8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   AFAACYTIN CHAIN RCM-BETA (FRAGMENT).
OS   Cerastes cerastes (Horned desert viper).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC   Viperidae; Viperinae; Cerastes.
RN   [1]
RP   SEQUENCE.
RX   MEDLINE; 96085138.
RA   Laraba-Djebari F., Martin-Eauclaire M.F., Mauco G., Marchot P.;
RT   "Afaacytin, an alpha beta-fibrinogenase from Cerastes cerastes (horned
RT   viper) venom, activates purified factor X and induces serotonin
RT   release from human blood platelets.";
RL   Eur. J. Biochem. 233:756-765(1995).
SQ   SEQUENCE   18 AA;  1937 MW;  CF680343EFE55327 CRC64;


  Query Match             21.2%;  Score 22;  DB 13;  Length 18;
  Best Local Similarity   36.4%;  Pred. No. 3.5e+03;
  Matches    4;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 NITGEKAMMLL 11
            ||   ::::||
Db      8 NINEHRSLVLL 18


RESULT  14
Q36277
ID   Q36277      PRELIMINARY;      PRT;    19 AA.
AC   Q36277;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   NADH DEHYDROGENASE SUBUNIT I (FRAGMENT).
GN   NAD1.
OS   Zea mays (Maize).
OG   Mitochondrion.
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; Liliopsida; Poales; Poaceae; Zea.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRB73; TISSUE=HYPOCOTYL;
RX   MEDLINE; 95140641.
RA   Thomson M.C., Macfarlane J.L., Beagley C.T., Wolstenholme D.R.;
RT   "RNA editing of mat-r transcripts in maize and soybean increases
RT   similarity of the encoded protein to fungal and bryophyte group II
RT   intron maturases: evidence that mat-r encodes a functional protein.";
RL   Nucleic Acids Res. 22:5745-5752(1994).
DR   EMBL; U09986; AAA67712.1; -.
DR   INTERPRO; IPR001694; -.
DR   PFAM; PF00146; NADHdh; 1.
KW   Mitochondrion.
FT   NON_TER       1      1
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2082 MW;  C2DA5548846188D3 CRC64;


  Query Match             21.2%;  Score 22;  DB 8;  Length 19;
  Best Local Similarity   62.5%;  Pred. No. 3.8e+03;
  Matches    5;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GEKAMMLL 11
          || | |:|
Db     11 GEYANMIL 18


RESULT  15
Q9S901
ID   Q9S901      PRELIMINARY;      PRT;    19 AA.
AC   Q9S901;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C6 PEPTIDE.
OS   Vigna sinensis (Cowpea).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Rosidae; eurosids I; Fabales; Fabaceae;
OC   Papilionoideae; Vigna.
RN   [1]
RP   SEQUENCE.
```

RA Hirano H., Kagawa H., Okubo K.;
RL Phytochemistry 31:731-735(1992).
SQ SEQUENCE 19 AA; 1989 MW; 6281AFF927726D74 CRC64;

Query Match 21.2%; Score 22; DB 10; Length 19;
Best Local Similarity 50.0%; Pred. No. 3.8e+03;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy       9 MLLGQVKYGL 18
           | :|||:  |
Db       1 MTVGQVQGNL 10
```

Search completed: December 21, 2000, 08:35:40
Job time: 603 sec